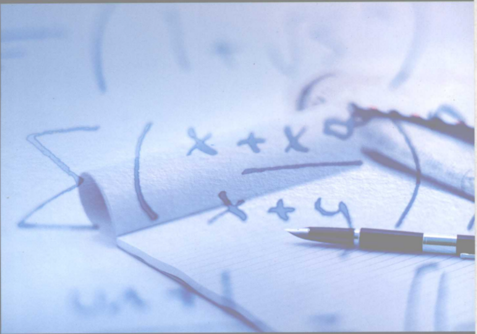
高等教育工科数学系列教材

现代数学基础



魏贵民 等
编著

XIANDAI SHUXUE JICHU

高等教育出版社
HIGHER EDUCATION PRESS

高等教育工科数学系列教材

微积分（上）

微积分（下）

线性代数

概率论与数理统计

离散数学

现代数学基础

ISBN 978-7-04-021765-0

9 787040 217650 >

定价：26.00元

现代数学基础

魏贵民　曹金文　魏友华
刘　锐　范安东　王权锋　余海洋　编著

高等教育出版社

内容提要

本书是高等教育工科数学系列教材之一，是一本基础性的现代数学教材，重点介绍与科学技术密切相关的一些重要的现代数学与应用数学的分支的基本概念和方法. 全书分七篇：第一篇 实变函数；第二篇 一般拓扑学；第三篇 泛函分析；第四篇 抽象代数；第五篇 人工神经网络；第六篇 小波变换；第七篇 分形理论及其应用. 前四篇介绍了现代数学的经典内容，后三篇介绍了与现代科技密切相关的一些现代应用数学内容及其在工程中的应用实例，以帮助读者更好地理解和掌握相关内容.

本书各篇可独立作为高校理工科类的高年级本科生和研究生的现代数学同名课程的教材或参考书，也可将各篇综合使用作为现代数学系统教材. 本书还可供科研工作者参考使用.

图书在版编目(CIP)数据

现代数学基础/魏贵民编著. —北京：高等教育出版社，2007.6

ISBN 978－7－04－021765－0

Ⅰ.现… Ⅱ.魏… Ⅲ.数学—高等学校—教材 Ⅳ.O1

中国版本图书馆 CIP 数据核字(2007)第 073043 号

策划编辑 徐 东 责任编辑 崔梅萍 封面设计 吴 昊 责任印制 蔡敏燕

出版发行	高等教育出版社	购书热线	010－58581118
社 址	北京市西城区德外大街4号		021－56964871
邮政编码	100011	免费咨询	800－810－0598
总 机	010－58581000	网 址	http://www.hep.edu.cn
传 真	021－56965341		http://www.hep.com.cn
			http://www.hepsh.com
		网上订购	http://www.landraco.com
经 销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
排 版	南京理工出版信息技术有限公司	畅想教育	http://www.widedu.com
印 刷	江苏如皋市印刷有限公司		
开 本	787×960 1/16	版 次	2007年6月第1版
印 张	22.5	印 次	2007年6月第1次
字 数	430 000	定 价	26.00元

物料号 21765－00

郑 重 声 明

前　言

经过长期的发展，现代数学不仅形成了一个分支众多的、庞大的、自身完美的理论知识体系，而且成为科学技术研究不可缺少的工具，它在众多的科学技术领域中的作用毋庸置疑. 随着当今计算机技术和其他的高新技术的发展以及数学核心领域的扩张，数学的空前广泛的应用以及计算机技术与数学更加紧密的结合与相互影响，使得现代数学方法在科技领域与国民经济领域发挥越来越重要的作用. 现代数学的研究与应用越来越受到高校科研工作者的关注与重视.

鉴于我国现行的数学教学中，全面系统地介绍现代数学基础理论知识的教材和参考书相对缺乏，特别是其中少有适用于理工科类的高年级本科生和研究生的书籍. 针对高校一般的理工科类学生的现状，本书编者们一致认为很有必要开设一门能够较系统、较全面地介绍现代数学内容的基础课程，以提高在校学生特别是研究生的现代数学的素质和培养创新思维的能力，增强他们运用数学知识分析解决实际问题的能力.

本书是一本概论性的现代数学教材，重点介绍与科学技术密切相关的一些重要的现代数学与应用数学的分支的基本概念和方法，为读者将来更加深入地学习和运用现代数学知识打下一定的基础. 本书由成都理工大学魏贵民教授与该校数学教学部、信息与计算科学系和应用数学系的主要专业骨干教师合著而成，他们在相应领域参与科学研究和教学工作多年，为本书的编写质量提供了有力的保障.

全书分七篇编写，主要包括：实变函数、一般拓扑学、泛函分析、抽象代数、人工神经网络、小波变换、分形理论及其应用，编写者为：魏贵民、曹金文、魏友华、刘锐、范安东、王权锋、余海洋. 前四篇从分析、几何、代数等的角度介绍了现代数学的基本内容，后三篇介绍了与现代科技密切相关的一些现代应用数学内容，同时各编者都给出了相应的理论在实际工程中的应用实例，以帮助读者更好地理解和掌握有关内容，这也是本书的特色之一. 本书涉及内容不但在数学科学中占

有相当重要的地位，而且在不同的科学领域中也有着广泛的应用.

在编写过程中，各编者取材广泛，论述严谨，重视基础，结构合理，深入浅出.该书可作为高校理工科类的高年级本科生和研究生的现代数学教材和参考书，也可供科研工作者参考使用.

限于编者水平，书中难免会出现一些错漏和不妥之处，还望有关专家和读者批评指正.

编　者

2007 年 3 月于成都理工大学

目　　录

第一篇　实 变 函 数

第二篇　一般拓扑学

第三篇　泛函分析

第四篇 抽象代数

第五篇　人工神经网络

第六篇　小 波 变 换

第七篇　分形理论及其应用

第一篇 实变函数

纯粹数学在 20 世纪得到了巨大的发展，产生出令人惊异的成就. 更高的抽象化是 20 世纪纯粹数学的主要趋势或特征之一. 集合论观点与公理化方法在 20 世纪逐渐成为数学抽象的范式，他们相互结合将数学的发展引向了高度抽象的道路，导致了 20 世纪上半叶实变函数论、泛函分析、拓扑学和抽象代数等具有标志性的四大抽象分支的崛兴. 这四大分支所创造的抽象语言、结构及方法，又渗透到数论、微分方程论、微分几何、代数几何、复变函数论及概率论等经典学科中，推动它们在抽象的基础上革新提高、演化发展.

现代数学是分支众多的、庞大的知识体系，大体说来，数学核心领域的扩张，数学的空前广泛的应用，以及计算机与数学的相互影响，形成了现代数学研究活动的三大方面.

集合论的观点在 20 世纪初首先引起了积分学的变革，从而导致了实变函数论的建立. 19 世纪末，分析的严格化迫使许多数学家认真考虑所谓“病态函数”，特别是不连续函数和不可微函数，并研究这样一个问题：积分的概念可以怎样推广到更广泛的函数类（如某种间断函数）上去. 这方面首先获得成功的是法国数学家 H. Lebesgue.

实变函数论是普通微积分的推广，它使微积分的适用范围大大扩展，引起数学分析的深刻变化，而 Lebesgue 积分正是这门学科的中心概念，作为分水岭，人们往往把 Lebesgue 以前的分析学叫做经典分析，而把以由 Lebesgue 积分引出的实变函数论为基础开拓出来的分析学叫做现代分析，因此 Lebesgue 积分可以看作是现代分析的开端，是 20 世纪数学的重大贡献之一.

第一章 集合的势

实变函数论是在集合论的观点与方法渗入数学分析的基础上产生的. 集合论的许多基本概念在实变函数论中都有着广泛的应用，本章将介绍一些有关集合论的基本知识，着重讲解集合的势等相关内容.

第一节　集合及其运算

集合是数学的基本概念之一.在以前的不少数学课程中读者对诸如集合、元素、子集、并集、交集、差集、补集、空集$\varnothing$、集合的表示方法等等内容已经非常熟悉,这里不再重述.

利用已知的集合可以诱导出一些新的集合,例如:

定义 1　集合 A 的所有子集为元素形成的集合

$$2^A = \{B \mid B \subseteq A\}$$

叫做集合 A 的幂集.

定义 2　已知集合 A 与集合 B,集合

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

叫做 A 与 B 的直积集.

我们把集合 $A \times B$ 的一个子集 φ 叫做从集合 A 到集合 B 的二元关系,而把从 B 到 A 的关系

$$\tilde{\varphi} = \{(b, a) \mid (a, b) \in \varphi\}$$

叫做 φ 的逆关系.

定义 3　非空集合 A 到 B 的关系 φ 如果满足: $\forall a \in A$, $\exists | b \in B$,使得 $(a, b) \in \varphi$,则 φ 叫做从集合 A 到集合 B 的映射,记为

$$\varphi: A \to B.$$

将 b 叫做 a 在 φ 下的像,记为 $b = \varphi(a)$,也可以表示为 $\varphi: a \mapsto b$. A 中所有元素在 φ 下的像的全体记为 $\varphi(A)$.

如果 $\varphi(A) = B$,就说 φ 是一个满射.

如果 A 中不同元素在 φ 下的像不同,就说 φ 是一个单射.

如果 φ 既是单射又是满射,就说 φ 是一个双射,也说 φ 是一个一一对应.双射 φ 的逆关系也是一个映射,叫做 φ 的逆映射,记为 φ^{-1}.

定义 4　设 $\varphi: A \to C$, $\psi: B \to C$ 是两个映射,如果 $A \subseteq B$,且对于 A 中每个元素 x 都有 $\varphi(x) = \psi(x)$,则 ψ 叫做 φ 在 B 上的延拓,记为 $\varphi \subseteq \psi$, φ 叫做 ψ 在 A 上的限制,记为 $\varphi = \psi|_A$.

例 1　设 $\varphi(x) = \sin x$, $x \in [0, \pi]$, $\psi(x) = |\sin x|$, $x \in (-\infty, +\infty)$,则 ψ 是 φ 在 $(-\infty, +\infty)$ 上的延拓,或 $\varphi = \psi|_{[0, \pi]}$.

关于集合的运算,除了读者已经熟知的集合的并集、交集、差集、补集外,在实变函数理论中还常用到集列的上限集和下限集.

定义 5 设 $A_1, A_2, \cdots, A_n, \cdots$ 是任意一列集合，由属于上述集列中无限多个集的那种元素全体所组成的集叫做这一集列的上限集或上极限，记为 $\overline{\lim\limits_{n\to\infty}} A_n$ 或 $\lim\limits_{n\to\infty} \sup A_n$，可表示为

$$\overline{\lim_{n\to\infty}} A_n = \{x \mid \text{存在无穷多个 } A_n \text{，使 } x \in A_n\}.$$

由上述集列中除去有限多个集外的所有集 A_n 都含有的那种元素所组成的集合叫做这一集列的下限集或下极限，记为 $\underline{\lim\limits_{n\to\infty}} A_n$ 或 $\lim\limits_{n\to\infty} \inf A_n$，可表示为

$$\underline{\lim_{n\to\infty}} A_n = \{x \mid \text{当 } n \text{ 充分大以后都有 } x \in A_n\}.$$

显然 $\underline{\lim\limits_{n\to\infty}} A_n \subseteq \overline{\lim\limits_{n\to\infty}} A_n$.

如果 $\underline{\lim\limits_{n\to\infty}} A_n = \overline{\lim\limits_{n\to\infty}} A_n$，则说集列 $\{A_n\}$ 收敛，此时 $A = \underline{\lim\limits_{n\to\infty}} A_n = \overline{\lim\limits_{n\to\infty}} A_n$ 叫做集列 $\{A_n\}$ 的极限集，记为 $A = \lim\limits_{n\to\infty} A_n$.

例 2 设 $A_n = \left[0, 1 + \dfrac{1}{n}\right]$，$(n = 1, 2, 3, \cdots)$，则因为

$$\underline{\lim_{n\to\infty}} A_n = \overline{\lim_{n\to\infty}} A_n = [0, 1],$$

故 $\lim\limits_{n\to\infty} A_n = [0, 1]$.

定理 1 对任意集列 $\{A_n\}$ 及集合 S，有

(1) $\overline{\lim\limits_{n\to\infty}} A_n = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m=n}^{\infty} A_m$，　$\underline{\lim\limits_{n\to\infty}} A_n = \bigcup\limits_{n=1}^{\infty} \bigcap\limits_{m=n}^{\infty} A_m$；

(2) $S - \underline{\lim\limits_{n\to\infty}} A_n = \overline{\lim\limits_{n\to\infty}} (S - A_n)$，　$S - \overline{\lim\limits_{n\to\infty}} A_n = \underline{\lim\limits_{n\to\infty}} (S - A_n)$.

证明留作习题.

定义 6 如果集列 $\{A_n\}$ 满足 $A_n \subseteq A_{n+1}$ $(A_n \supseteq A_{n+1})$，$n = 1, 2, 3, \cdots$，则 $\{A_n\}$ 叫做单调增加（单调减少）集列. 单调增加与单调减少集列统称为单调集列.

显然有：

定理 2 单调集列是收敛的，并且如果 $\{A_n\}$ 单调增加，则 $\lim\limits_{n\to\infty} A_n = \bigcup\limits_{n=1}^{\infty} A_n$，如果 $\{A_n\}$ 单调减少，则 $\lim\limits_{n\to\infty} A_n = \bigcap\limits_{n=1}^{\infty} A_n$.

例 3 设 $A_n = \left[0, 1 + \dfrac{1}{n}\right)$，证明 $\bigcap\limits_{n=1}^{\infty} A_n = [0, 1]$.

证 因为 $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq A_{n+1} \supseteq \cdots$，所以 $\{A_n\}$ 是单调减少集列，故 $\bigcap\limits_{n=1}^{\infty} A_n = \lim\limits_{n\to\infty} A_n = \underline{\lim\limits_{n\to\infty}} A_n = [0, 1]$.

最后，介绍集合的特征函数.

定义 7 设 X 是非空集合，又 $A \subseteq X$，作 X 上的函数

$$\chi_A(x) = \begin{cases} 1, & \text{当 } x \in A \text{ 时}; \\ 0, & \text{当 } x \notin A \text{ 时}. \end{cases}$$

则 $\chi_A(x)$ 叫做集合 A 的特征函数.

定理 3 (特征函数的性质)设 X 是非空集合，A, B, $A_n(n = 1, 2, \cdots)$ 都是 X 的子集，则有：

(1) $A = X$ 的充分必要条件是 $\chi_A(x) = 1$,

$A = \varnothing$ 的充分必要条件是 $\chi_A(x) = 0$;

(2) $A \subseteq B$ 的充分必要条件是 $\chi_A(x) \leqslant \chi_B(x)$,

$A = B$ 的充分必要条件是 $\chi_A(x) = \chi_B(x)$;

(3) $\chi_{\bigcup\limits_{n=1}^{\infty} A_n}(x) = \max \chi_{A_n}(x)$, $\chi_{\bigcap\limits_{n=1}^{\infty} A_n}(x) = \min \chi_{A_n}(x)$;

(4) 极限集 $\lim\limits_{n\to\infty} A_n$ 存在的充分必要条件是极限 $\lim\limits_{n\to\infty} \chi_{A_n}(x)$ 存在，而且此时必有 $\chi_{\lim\limits_{n\to\infty} A_n}(x) = \lim\limits_{n\to\infty} \chi_{A_n}(x)$.

第二节　集合的势

比较两个集合所含元素的多少，对于有限集，用数数的办法，对于两个无限集，这样的办法就不行了. 下面将利用一一对应的思想方法解决这个问题.

定义 1 设 A 和 B 是非空集合，如果存在一一对应 $\varphi: A \to B$，则说 A 和 B 对等，记为 $A \sim B$. 规定 $\varnothing \sim \varnothing$.

例 1 奇数集，偶数集，自然数集是彼此对等的.

例 2 $(0, 1) \sim (0, +\infty)$.

事实上，只要令 $\varphi(x) = \dfrac{x}{1-x}$, $x \in (0, 1)$ 即可.

例 2 说明无限集可以与它的某一个真子集对等，可以证明：**任意一个无限集都能与它的一个真子集对等**，而有限集是不能与其真子集对等的. 这说明了无限集和有限集的区别. 这可以用来作为有限集与无限集的定义.

定义 2 能与其一个真子集对等的非空集合叫做无限集，否则，叫做有限集.

由定义可知对等关系显然具有下列性质：

定理 1 对于集合 A, B, C，有

(1) 反身性：　$A \sim A$;

(2) 对称性：　若 $A \sim B$，则 $B \sim A$;

(3) 传递性：　若 $A \sim B$, $B \sim C$，则 $A \sim C$.

定理 2 设 $\{A_\alpha \mid \alpha \in \Gamma\}$, $\{B_\alpha \mid \alpha \in \Gamma\}$ 是两集族，若 $A_\alpha \sim B_\alpha$ 且 $A_\alpha \cap A_\beta = \varnothing$, $B_\alpha \cap B_\beta = \varnothing$ $(\alpha \neq \beta,\ \alpha,\ \beta \in \Gamma)$，则有 $\bigcup_{\alpha\in\Gamma} A_\alpha \sim \bigcup_{\alpha\in\Gamma} B_\alpha$.

证 由于 $A_\alpha \sim B_\alpha(\alpha \in \Gamma)$，故存在一一对应 $f_\alpha: A_\alpha \to B_\alpha$，作映射 $f: \bigcup_{\alpha\in\Gamma} A_\alpha \to \bigcup_{\alpha\in\Gamma} B_\alpha$ 使得 $f(x) = f_\alpha(x)$ $(x \in A_\alpha,\ \alpha \in \Gamma)$，显然 f 是一一对应.

定义 3 设 A, B 是两个集合，如果 $A \sim B$，则说 A 和 B 具有相同的势（或基数）.

记集合 A 的势为 $|A|$，则 $|A| = |B|$ 当且仅当 $A \sim B$.

定义 4 如果 A 与 B 不对等，但存在 B 的真子集 B^*，有 $A \sim B^*$，则说 A 的势小于 B 的势（或 B 的势大于 A 的势），记为 $|A| < |B|$（或 $|B| > |A|$）.

从逻辑上讲，还有可能集合 A 的真子集 A^* 与集合 B 对等，而集合 B 也有真子集 B^* 与 A 对等. 而人们证明了

定理 3 （F. Bernstein 定理）设 A, B 是两个集合，若有 A 的子集 A^* 以及 B 的子集 B^*，使 $A \sim B^*$, $B \sim A^*$，则 $A \sim B$.

定理 3 的证明较繁杂，这里略去，读者可参看本篇末所列的参考文献.

利用势的说法，定理 3 可叙述为：若 $|A| \leqslant |B|$, $|B| \leqslant |A|$，则 $|A| = |B|$.

势的概念可以看作有限集合中所含元素个数的推广. 事实上，若 A 为一非空有限集，则 A 必能与 $\{1, 2, 3, \cdots\}$ 中的某一个截段 $\{1, 2, \cdots, n\}$ 一一对应，而 $\{1, 2, \cdots, n\}$ 的势可以简单地用 n 来表示，故有 $|A| = n$. 而对无限集合来说，$|A|$ 当然就是通常元素个数概念的推广.

推论 1 设 $A \subseteq B \subseteq C$，若 $A \sim C$，则 $B \sim C$.

推论 2 设 A, B 是两个集合，则关系式

$$|A| < |B|,\quad |A| = |B|,\quad |A| > |B|$$

有且只有一个成立.

第三节 可数集合

一、可数集合的概念和性质

定义 1 与自然数集 **N** 对等的集合叫做可数集合.

可数集合的势记为 $\aleph_0$.

若 A 是有限集或可数集，就说 A 是至多可数集.

因为自然数集 $\{0, 1, 2, 3, \cdots\}$ 的元素可以排成无穷序列的形式，因此显然有下列事实：

定理 1 集合 A 是可数集的充分必要条件是 A 的元素可排成无穷序列的

形式

$$A=\{a_0,a_1,a_2,\cdots,a_n,\cdots\}.$$

定理 2　任意无限集合都至少包含一个可数子集.

证　设 M 是一个无限集合,因 $M\neq\varnothing$,故可任取一个元素 $e_0\in M$,且 $M-\{e_0\}\neq\varnothing$,于是又可任取一个元素 $e_1\in M-\{e_0\}$,显然 $e_1\in M$,且 $e_0\neq e_1$. 设已经取出 n 个互不相同的元素 $e_0,e_1,e_2,\cdots,e_{n-1}(e_i\in M,i=0,1,2,\cdots,n-1)$,由于 $M-\{e_0,e_1,e_2,\cdots,e_{n-1}\}\neq\varnothing$,于是可任取一元素 $e_n\in M-\{e_0,e_1,e_2,\cdots,e_{n-1}\}$,显然 $e_n\in M$ 且 $e_n\neq e_i(i=0,1,2,\cdots,n-1)$. 由归纳法,可得 M 的一个无限子集 $\{e_0,e_1,e_2,\cdots,e_n,\cdots\}$,它显然为一可数集.

注意　该定理说明**可数集在所有无限集中具有最小的势.**

定理 3　可数集合的子集或者是有限集或者是可数集.

证　设 A 为可数集,$A^*\subseteq A$,若 A^* 是无限集合,则必有

$$|A^*|\leqslant|A|=\aleph_0,$$

又因为 A^* 为无限集,$|A^*|\geqslant\aleph_0$,从而 $|A^*|=\aleph_0$,即 A^* 是可数集.

定理 4　有限个或可数个至多可数集的并集至多可数.

证　不失一般性,设有一列集合 $A_i(i=0,1,2,\cdots,n$ 或 $i=0,1,2,\cdots)$ 均为可数集. 于是可令

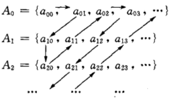


图 1.1.1

将 $\bigcup\limits_{i=0}^{n}A_i$ 或 $\bigcup\limits_{i=0}^{\infty}A_i$ 中元素按上图箭头所指顺序排成一列

$$a_{00},a_{01},a_{10},a_{20},a_{11},a_{02},a_{03},a_{12},a_{21},a_{30},\cdots,$$

若集列 $\{A_i\}$ 中某两集合有公共元素,因为公共元素在并集中只算一次,则在序列中去掉重复的元素后,所得集合为 $\bigcup\limits_{i=0}^{n}A_i$ 或 $\bigcup\limits_{i=0}^{\infty}A_i$,仍是可数的.

对于 $A_0,A_1,A_2,\cdots$ 中有些是有限集,或仅为有限个集的情况,也可类似地讨论.

定理 5　若 A 中每个元素由 k 个互相独立的记号所决定,各记号跑遍一个可数集,即 $A=\{a_{x_1x_2\cdots x_k}\mid x_i=x_i^{(0)}x_i^{(1)}\cdots,i=1,2,\cdots,k\}$,则 A 是可数集.

证　用数学归纳法证明.

(1) $k=1$ 时，结论显然成立.

(2) 假设 $k=n$ 时结论成立，即 $\{a_{x_1x_2\cdots x_n} \mid x_i = x_i^{(0)}x_i^{(1)}\cdots,\ i=1,\ 2,\ \cdots,\ n\}$ 可数.

下证 $k=n+1$ 时，结论成立：

令 $A_j = \{a_{x_1x_2\cdots x_nx_{n+1}^{(j)}} \mid x_i = x_i^{(0)}x_i^{(1)}\cdots,\ i=1,\ 2,\ \cdots,\ n\}$，则由假设 A_j 可数，故 $\bigcup\limits_{j=0}^{\infty} A_j = \{a_{x_1x_2\cdots x_nx_{n+1}} \mid x_i = x_i^{(0)}x_i^{(1)}\cdots,\ i=1,\ 2,\ \cdots,\ n,\ n+1\}$ 也可数.

综上可知结论成立.

二、可数集合的例

例 1 自然数集 $\mathbf{N}$，正整数集 $\mathbf{Z}^+$，整数集 $\mathbf{Z}$，均为可数集.

例 2 三角函数系 $\{1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos nx,\ \sin nx,\ \cdots\}$ 是可数集.

例 3 全体有理数之集 $\mathbf{Q}$ 是可数集.

证 用 $\mathbf{Q}^+$，$\mathbf{Q}^-$ 分别表示正，负有理数集，则 $\mathbf{Q} = \mathbf{Q}^+ \cup \mathbf{Q}^- \cup \{0\}$.

设 $A_i = \left\{\dfrac{1}{i},\ \dfrac{2}{i},\ \dfrac{3}{i},\ \cdots\right\}$ $(i=1,\ 2,\ 3,\ \cdots)$，则 A_i 是可数集，于是 $\mathbf{Q}^+ = \bigcup\limits_{i=1}^{\infty} A_i$ 可数.

同理可证 $\mathbf{Q}^-$ 可数.

故由定理 4 可知 $\mathbf{Q} = \mathbf{Q}^+ \cup \mathbf{Q}^- \cup \{0\}$ 可数.

有理数在实数中是处处稠密的，可它却是一个和数轴上稀疏分布着的自然数全体有着一一对应的可数集，这个表面看来令人难以置信的事实，充分说明了要判断一个集合是否可数应该细致分析判定.

例 4 平面上坐标为有理数的点所成集合 A 是可数集合.

证 因为 $A = \{(a,\ b) \mid a,\ b$ 是有理数$\}$，由定理 5，可知 A 可数.

例 5 代数数全体成一可数集(所谓代数数，是整系数多项式的根).

证 设整系数多项式 $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$，对固定的 n，由定理 5 可知，整系数 n 次多项式全体为一可数集. 再用定理 4，可知整系数多项式全体为一可数集. 而每个多项式只有有限个根，于是再利用定理 4，可知结论成立.

第四节　不可数集合

不是可数集的无限集合叫做不可数集合. 本节介绍常见的不可数集合.

定理 1 **区间$[0,\ 1]$是一个不可数集合.**

证 假设$[0,\ 1]$是一个可数集，则$[0,\ 1]$中点可排成一个无穷序列

$$x_1, x_2, \cdots, x_n, \cdots$$

的形式. 记 $[0, 1] = I_0$,作闭区间 $I_1 \subset I_0$,使 I_1 的长度 $|I_1| < \frac{1}{2}$,且 $x_1 \notin I_1$. 再作闭区间 $I_2 \subset I_1$,使 $|I_2| < \frac{1}{2^2}$,且 $x_2 \notin I_2$,显然此时 $x_1 \notin I_2$. 设已作出符合要求的闭区间 $I_1, I_2, \cdots, I_n$,于是有 $I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_n$,且 $|I_i| < \frac{1}{2^i}$ $(i = 1, 2, \cdots, n)$. 再作闭区间 $I_{n+1} \subset I_n$,使 $|I_{n+1}| < \frac{1}{2^{n+1}}$,且 $x_{n+1} \notin I_{n+1}$. 这样一来,可得一个闭区间列

$$I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots,\ |I_n| < \frac{1}{2^n},\ x_n \notin I_n\ (n = 1, 2, \cdots),$$

因为 $\frac{1}{2^n} \to 0(n \to \infty)$,由数学分析中闭区间套定理知,存在惟一一点 $x_0 \in I_n$ $(n = 1, 2, \cdots)$,而 $I_n \subset I_0$,故 $x_0 \in I_0 = [0, 1]$. 但 x_0 不在序列中(因为如果 $x_0 = x_n$ $(n \geqslant 1)$,则 $x_0 \notin I_n$,与 x_0 的定义产生矛盾),这与[0, 1]中的点可排成序列 $x_1, x_2, \cdots, x_n, \cdots$矛盾,故假设不成立,即[0, 1]是不可数集.

定义 与区间[0, 1]对等的集合叫做连续统.

连续统的势叫做连续势,记为 $\aleph_1$.

推论 1 $\aleph_1 > \aleph_0$.

证 由定理 1 可知 $\aleph_1 \neq \aleph_0$,而 $[0, 1] \supset \left\{1, \frac{1}{2}, \frac{1}{3}, \cdots\right\} \sim \{1, 2, 3, \cdots\}$,故 $\aleph_1 > \aleph_0$.

例 1 区间(0, 1)具有连续势.

证 将(0, 1)中的全体有理数排列为

$$r_1, r_2, \cdots, r_n, \cdots,$$

而[0, 1]中的全体有理数可排列为

$$0, 1, r_1, r_2, \cdots, r_n, \cdots,$$

作映射 f 如下:

当 $x = r_1$ 时,$f(x) = 0$;

当 $x = r_2$ 时,$f(x) = 1$;

当 $x = r_n$ 时,$f(x) = r_{n-2}$ $(n > 2)$;

当 x 是(0, 1)中的无理数时,$f(x) = x$.

则 $f(x)$是(0, 1)与[0, 1]间的一一对应.

例 2 全体实数所成之集的势为$\aleph_1$.

证 作$(0,1)$与$(-\infty, +\infty)$之间的一一对应 $f(x)=\tan\left(\pi x-\frac{\pi}{2}\right)$即可.

例 3 任意区间(a, b)、$[a, b)$、$(a, b]$、$[a, b]$、$(a, +\infty)$、$[a, +\infty)$、$(-\infty, b)$、$(-\infty, b]$均具有连续势$\aleph_1$.

例 4 实数列全体 $\mathbf{E}_\infty$ 的势为$\aleph_1$.

证 记$B=\{(x_1, x_2, \cdots, x_n, \cdots) \mid 0<x_n<1, n=1, 2, \cdots\}$,则$B\subset \mathbf{E}_\infty$.
作映射

$$\varphi:(x_1, x_2, \cdots, x_n, \cdots) \mapsto \left\{\tan\left(x_1-\frac{1}{2}\right)\pi,\right.$$

$$\left.\tan\left(x_2-\frac{1}{2}\right)\pi, \cdots, \tan\left(x_n-\frac{1}{2}\right)\pi, \cdots\right\}.$$

显然 $B \overset{\varphi}{\sim} \mathbf{E}_\infty$.

下面证明 $|B|=\aleph_1$.

易知$(0,1)$对等于 B 的一个子集 $B_0=\{(x, x, \cdots, x, \cdots) \mid 0<x<1\}$,反过来,对于 B 中任意的 $x=(x_1, x_2, \cdots, x_n, \cdots)$,设

$$x_i=0.\ x_{i1}x_{i2}x_{i3}x_{i4}\cdots\ (i=1, 2, \cdots),$$

作映射

$$\psi: x \mapsto 0.\ x_{11}x_{12}x_{21}x_{31}x_{22}x_{13}x_{14}x_{23}\cdots,$$

其中 x_{ij} 的排列顺序按下图中箭头所示:

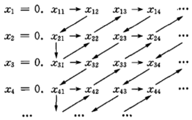


图 1.1.2

显然 $\psi(x)\in(0, 1)$,且易验证 ψ 是 B 到$(0,1)$的一个子集上的一一对应.
故由 F. Bernstein 定理可得 $B\sim(0, 1)$,所以 $|B|=\aleph_1$.

例 5 集合 $\mathbf{R}^n=\{(x_1, x_2, \cdots, x_n) \mid x_i\in(-\infty, +\infty), i=1, 2, \cdots, n\}$ 的势是 $\aleph_1$.

证 将 $\mathbf{R}^n$ 中点$(x_1, x_2, \cdots, x_n)$对应于 $\mathbf{E}_\infty$ 中的点$(x_1, x_2, \cdots, x_n, 0, \cdots, 0, \cdots)$时,就知道 $\mathbf{R}^n$ 对等于 $\mathbf{R}^1$ 的子集. 如果再将 $\mathbf{R}^1$ 中点 x 对应 $\mathbf{R}^n$ 中点$(x,$

0，…，0)时，又知道 $\mathbf{R}^1$ 对等于 $\mathbf{R}^n$ 的一个子集. 所以由 F. Bernstein 定理知道 $|\mathbf{R}^n| = |\mathbf{R}^1| = \aleph_1$.

定理 2 **设 $A_1, A_2, \cdots, A_n, \cdots$ 是一列互不相交的集合，它们的势都是 $\aleph_1$，则 $\bigcup_{n=1}^{\infty} A_n$ 的势也是 $\aleph_1$.**

证 设 $I_n = (n-1, n]$，则 $I_n \cap I_m = \varnothing\ (n \neq m)$，但 $|I_n| = \aleph_1\ (n = 1, 2, \cdots)$，故 $I_n \sim A_n\ (n = 1, 2, \cdots)$. 从而 $\bigcup_{n=1}^{\infty} A_n \sim \bigcup_{n=1}^{\infty} I_n = (0, +\infty)$.

定理 3 **$\aleph_1$ 个势为 $\aleph_1$ 的集的并集的势为 $\aleph_1$.**

事实上，对于每一个被并的集合 A，可让 A 与 xOy 平面上平行于 Ox 轴的直线上全体点的集合一一对应，于是可得所述并集与 xOy 平面上全体点的集合之间一一对应.

定理 4 **设 M 是任意一个集合，它的所有子集形成集合即 M 的幂集 2^M，有 $|2^M| > |M|$.**

证 显然 M 对等于 2^M 的一个子集，因为映射 $\varphi: x \mapsto \{x\}$（其中 $x \in M$，$\{x\} \in 2^M$）是 M 到 2^M 的一个子集的一一对应.

而 2^M 一定不能对等于 M. 如若不然，即 $2^M \sim M$，则 $\forall x \in M$，都应有 M 的一个子集 M_x 与之对应. 将 M 中所有满足 $x \notin M_x$ 的元素 x 放在一起作成集合 M^1，则 $M^1 \subseteq M$，故 $M^1 \in 2^M$. 由于 2^M 和 M 对等，所以应该有 M 中元素 x^1 与 M^1 对应，由 M^1 的定义可知 $x^1 \notin M^1$ 且 $x^1 \in M^1$，这就产生了矛盾，因而 2^M 不对等于 M.

于是就有 $|2^M| > |M|$.

定理 4 告诉我们：**没有一个最大的势.**

习 题 1-1

1. 证明：对任意集列 $\{A_n\}$ 及集合 S，有

(1) $\overline{\lim_{n\to\infty}} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$，$\underline{\lim_{n\to\infty}} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$；

(2) $S - \underline{\lim_{n\to\infty}} A_n = \overline{\lim_{n\to\infty}}(S - A_n)$，$S - \overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}}(S - A_n)$.

2. 设 $A_{2n-1} = \left(0, \frac{1}{n}\right)$，$A_{2n} = (0, n)\ (n = 1, 2, \cdots)$，求集列 $\{A_n\}$ 的上限集和下限集.

3. 作出一个 $(-1, 1)$ 和 $(-\infty, +\infty)$ 的一一对应.

4. 证明：将球面去掉一点后，余下的点所成的集合与整个平面上的点所成的集合是对等的.

5. 证明：由直线上互不相交的开区间为元素的集合 A 是至多可数集.

6. 设集合 A 是平面上以有理点为中心，有理数为半径的圆的全体，则 A 是可数集.

7. 证明:单调增加函数的不连续点最多只有可数个.

8. 证明:[0, 1]上的全体无理数所成之集具有连续势$\aleph_1$.

9. 如果$A=\bigcup\limits_{n=1}^{\infty}A_n$,且$|A|=\aleph_1$,则至少存在一个$A_n$,使得$|A_n|=\aleph_1$.

第二章　欧氏空间中的点集

本章介绍 n 维 Euclid(欧几里得)空间中的点集理论,它们是数学分析课程中几何空间的点集理论的推广,为讨论点集的测度问题做准备.

第一节　欧 氏 空 间

设实数集 $\mathbf{R}$. 集合

$$\mathbf{R}^n = \{x \mid x = (x_1, x_2, \cdots, x_n),\ x_i \in \mathbf{R},\ i = 1, 2, \cdots, n\},$$

$$x = (x_1, x_2, \cdots, x_n),\ y = (y_1, y_2, \cdots, y_n)$$

是 $\mathbf{R}^n$ 的两个点.

按照加法

$$x + y = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

及数量乘法

$$kx = (kx_1, kx_2, \cdots, kx_n) \qquad (k \in \mathbf{R})$$

便构成n 维线性空间. 如果定义 x, y 的内积

$$x \cdot y = \sum_{i=1}^{n} x_i y_i,$$

则 $\mathbf{R}^n$ 便构成一个n 维欧氏空间. 为了强调这个 n 维欧氏空间与集合 $\mathbf{R}^n$ 的区别,我们常将这个 n 维欧氏空间记为 $\mathbf{E}^n$,而 $\mathbf{R}^n$ 只是 $\mathbf{E}^n$ 的点的集合.

在 n 维欧氏空间 $\mathbf{E}^n$ 中,两点 x, y 的距离是

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

当 $n = 1, 2, 3$ 时,n 维欧氏空间 $\mathbf{E}^n$ 便是直线、平面以及我们生活的物理空间中的几何空间. 因此,n 维欧氏空间是直观的几何空间的自然的推广.

显然, $\forall x, y, z \in \mathbf{E}^n$,有

(1) **正则性:**　$d(x, y) \geqslant 0$, $d(x, y) = 0$ 当且仅当 $x = y$;

(2) **对称性:**　$d(x, y) = d(y, x)$;

(3) **三角不等式:**　$d(x, y) \leqslant d(x, z) + d(z, y)$.

由两点间的距离可定义 $\mathbf{E}^n$ 中非空点集 E 的直径为

$$\delta(E)=\sup_{x\in E,y\in E}d(x,y),$$

非空点集 A，B 之间的距离为

$$d(A,B)=\inf_{x\in A,y\in B}d(x,y).$$

如果 $\delta(E)<+\infty$，则说 E 是 $\mathbf{E}^n$ 中的有界点集.

规定空集也是有界点集.

设 $x_0\in\mathbf{E}^n$，$\delta>0$，点集 $\{x\mid d(x_0,x)<\delta,x\in\mathbf{E}^n\}$ 叫做点 x_0 的 δ 邻域，记为 $O(x_0,\delta)$. $\{x_k\}$ $(k=1,2,\cdots)$ 是 $\mathbf{E}^n$ 中一点列，如果 $\forall\varepsilon>0$，$\exists k_0\in\mathbf{Z}^+$，当 $k>k_0$ 时有 $x_k\in O(x_0,\varepsilon)$，则说 $\{x_k\}$ 收敛于 x_0，记为

$$\lim_{k\to\infty}x_k=x_0 \text{ 或 } x_k\to x_0(k\to\infty).$$

点集 $\{(x_1,x_2,\cdots,x_n)\mid a_i<x_i<b_i,i=1,2,\cdots,n\}$ 叫做一个 n 维开区间. 若其中不等式换成 $a_i\leqslant x_i\leqslant b_i$ 或 $a_i<x_i\leqslant b_i$ 或 $a_i\leqslant x_i<b_i$，$i=1,2,\cdots,n$，则叫做一个 n 维闭区间或 n 维左开右闭区间或 n 维左闭右开区间，无需区别时，统称为区间，记为 I. b_i-a_i $(i=1,2,\cdots,n)$ 叫做 I 的第 i 个"边长". $\prod_{i=1}^{n}(b_i-a_i)$ 叫做 I 的"体积"，记为 $|I|$.

规定空集的体积为 0.

第二节　内点　聚点　边界点

设 E 是 $\mathbf{E}^n$ 中的一个点集，x_0 是 $\mathbf{E}^n$ 中的一定点，则 x_0 与 E 的关系有如下三种可能：

(1) 在 x_0 附近全是 E 的点；

(2) 在 x_0 附近没有 E 的点；

(3) x_0 附近既有 E 的点，又有不属于 E 的点.

针对这三种情形，我们有下述定义：

定义 1　如果存在 $O(x_0,\delta)\subseteq E$，则 x_0 叫做 E 的内点. E 的全体内点作成的集合叫做 E 的内部，记为 E^0 或 $\mathrm{Int}E$.

如果任意 $O(x_0,\delta)$ 内都有 E 的无穷多个点，则 x_0 叫做 E 的聚点. E 的全体聚点作成的集合叫做 E 的导集，记为 E'. $E\cup E'$ 叫做 E 的闭包，记为 $\overline{E}$.

如果任意 $O(x_0,\delta)$ 内既有属于 E 的点，也有不属于 E 的点，则 x_0 叫做 E 的边界点. E 的全体边界点作成的集合叫做 E 的边界，记为 E^b.

如果存在 $O(x_0, \delta)$，使得 $E \cap O(x_0, \delta) = \{x_0\}$. 则 x_0 叫做 E 的孤立点.

显然，E 的内点必为聚点，但反之不然；而 E 的边界点不是聚点便是孤立点. 其次，E 的内点必属于 E，但 E 的聚点可属于 E 也可不属于 E.

定理 1 **下面三个命题等价：**

(1) x_0 是 E 的聚点；

(2) x_0 的任一邻域内，至少含有一个属于 E 而异于 x_0 的点；

(3) 存在 E 中互异的点所成点列 $\{x_n\}$，使得 $x_n \to x_0 (n \to \infty)$.

证 由聚点的定义可知(1)推出(2)，(3)推出(1)是显然的.

下面证明(2)推出(3)：

由已知，存在 $O(x_0, \delta)$，其中至少有一个点 $x_1 \in E$，且 $x_1 \neq x_0$.

令 $\delta_1 = \min\left\{d(x_1, x_0), \dfrac{\delta}{2}\right\}$，则在 $O(x_0, \delta_1)$ 中至少有一点 $x_2 \in E$ 且 $x_2 \neq x_0$.

令 $\delta_2 = \min\left\{d(x_2, x_0), \dfrac{\delta}{3}\right\}$，则在 $O(x_0, \delta_2)$ 中至少有一点 $x_3 \in E$ 且 $x_3 \neq x_0$.

这样继续下去，可得点列 $\{x_n\}$，满足结论要求.

定理 2 (1) $(E^0)^c = \overline{E^c}$, $(\overline{E})^c = (E^c)^0$ （这里 E 的补集 E^c 是对 $\mathbf{E}^n$ 而言的，以后仿此）；

(2) 若 $A \subseteq B$，则 $A' \subseteq B'$, $A^0 \subseteq B^0$, $\overline{A} \subseteq \overline{B}$；

(3) $(A \cup B)' = A' \cup B'$, $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

证 (1)(2)可由定义直接推出.

下面证明(3)：

因为 $A \subseteq A \cup B$, $B \subseteq A \cup B$，由(2)可知，$A' \subseteq (A \cup B)'$, $B' \subseteq (A \cup B)'$，从而 $A' \cup B' \subseteq (A \cup B)'$.

另一方面，设 $x_0 \in (A \cup B)'$，则由定理 1 可知有一互异点列 $\{x_n\}$，$x_n \in A \cup B$，使得 $d(x_n, x_0) \to 0 (n \to \infty)$.

若 $x_0 \in A'$，则 $x_0 \in A' \cup B'$；

若 $x_0 \notin A'$，则 $\{x_n\}$ 中最多只有有限个点属于 A，其余无穷多个点全属于 B. 于是由定理 1 可知 $x_0 \in B'$，从而 $x_0 \in A' \cup B'$.

总之，$x_0 \in A' \cup B'$.

于是综合两方面可得 $(A \cup B)' = A' \cup B'$.

而 $\overline{A \cup B} = (A \cup B) \cup (A \cup B)' = (A \cup B) \cup (A' \cup B')$

$= (A \cup A') \cup (B \cup B') = \overline{A} \cup \overline{B}$.

定理 3 (Bolzano-Weierstrass **定理**) **$\mathbf{E}^n$ 中的有界点列必有收敛子列.**

证 设 $\{x^{(k)}\} = \{(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})\}$ 是 $\mathbf{E}^n$ 中的任一有界点列，则每个 $\{x_i^{(k)}\}$ $(i = 1, 2, \cdots, n)$ 都是有界数列. 由 $\mathbf{E}^1$ 中的 Bolzano-Weierstrass 定理可

知$\{x_1^{(k)}\}$有收敛子列$\{x_1^{(k_1)}\}$，其中$\{k_1\}$是正整数列$\{k\}$的子列.

作$\{x^{(k)}\}$的子列

$$\{x^{(k_1)}\} = \{(x_1^{(k_1)},\ x_2^{(k_1)},\ \cdots,\ x_n^{(k_1)})\},$$

其中$\{x_1^{(k_1)}\}$是收敛数列.

同理，由于$\{x_2^{(k_1)}\}$是有界数列，所以$\{x^{(k_1)}\}$有子列

$$\{x^{(k_2)}\} = \{(x_1^{(k_2)},\ x_2^{(k_2)},\ \cdots,\ x_n^{(k_2)})\},$$

其中$\{x_2^{(k_2)}\}$是收敛数列. 此时，由于$\{x_1^{(k_2)}\}$是$\{x_1^{(k_1)}\}$的子列，从而$\{x_1^{(k_2)}\}$也是收敛数列. 如此一直进行到第 n 步，就得到$\{x^{(k)}\}$的子列

$$\{x^{(k_n)}\} = \{(x_1^{(k_n)},\ x_2^{(k_n)},\ \cdots,\ x_n^{(k_n)})\},$$

其中 $\{x_i^{(k_n)}\}$ $(i = 1,\ 2,\ \cdots,\ n)$ 都是收敛数列. 设$\{x_i^{(k_n)}\}$ 收敛于 $x_i^{(0)}$ $(i = 1,\ 2,\ \cdots,\ n)$，记$x^{(0)} = (x_1^{(0)},\ x_2^{(0)},\ \cdots,\ x_n^{(0)})$，由不等式

$$d(x^{(k_n)},\ x^{(0)}) = \Big(\sum_{i=1}^{n} |\ x_i^{(k_n)} - x_i^{(0)}\ |^2\Big)^{\frac{1}{2}} \leqslant \sqrt{n}\max_{1\leqslant i\leqslant n}|\,x_i^{(k_n)} - x_i^{(0)}\,|$$

知$\{x^{(k_n)}\}$收敛于 $x^{(0)}$.

推论　$\mathbf{E}^n$ 中的有界无限点集至少有一个聚点.

证　设 E 是 $\mathbf{E}^n$ 中的任一个有界无限点集，可从 E 中取出一个各项互异的有界点列$\{x_k\}$，由上述定理可知$\{x_k\}$有子列$\{x_{k_i}\}$收敛于某点 x_0，由于$\{x_{k_i}\}$也是 E 中各项互异点列，因此 x_0 是 E 的聚点.

第三节　开集　闭集　完备集

本节重点介绍 $\mathbf{E}^n$ 中两类特殊点集，并对其性质进行一些讨论.

定义 1　**若 $E \subseteq E^0$，则 E 叫做开集.**

例如，在 $\mathbf{E}^1$ 中$(a,\ b)$是开集，整个空间 $\mathbf{E}^n$ 是开集，空集也是开集.

定义 2　**若 $E' \subseteq E$，则 E 叫做闭集.**

例如，在 $\mathbf{E}^1$ 中$[a,\ b]$是闭集，整个空间 $\mathbf{E}^n$ 是闭集，空集也是闭集.

容易证明：**E 为开集的充分必要条件是 $E^0 = E$，E 为闭集的充分必要条件是 $\overline{E} = E$.**

定理 1　**对任意 $E \subseteq \mathbf{E}^n$，E^0 是开集，E'和 $\overline{E}$ 是闭集.**

证　设 $x_0 \in E^0$，由 E^0 的定义可知存在 $O(x_0,\ \delta) \subseteq E$. 任取 $x_1 \in O(x_0,\ \delta)$，可知存在 $O(x_1,\ \delta_1) \subseteq O(x_0,\ \delta) \subseteq E$，于是 $x_1 \in E^0$，所以 $O(x_0,\ \delta) \subseteq E^0$，即 x_0 是 E^0 的内点，由开集的定义可知 E^0 为开集.

其次证明 E' 为闭集：

设 $x_0 \in (E')'$，由第二节定理 1 中的(2)可知，在 x_0 的任意邻域 $O(x_0, \delta)$ 中必然有点 $x_1 \in E'$，且 $x_1 \neq x_0$. 因为 $x_1 \in E'$，于是又有属于 E 的点 $x_2 \in O(x_0, \delta)$，且 $x_2 \neq x_0$. 于是由第二节定理 1 可知，$x_0 \in E'$. 所以 $(E')' \subseteq E'$，即 E' 是闭集.

最后证明 $\overline{E}$ 是闭集：

因为 $\overline{E} = E \cup E'$，所以 $(\overline{E})' = E' \cup (E')' \subseteq E' \cup E' = E' \subseteq \overline{E}$，因此 $\overline{E}$ 是闭集.

定理 2 （开集与闭集的对偶性）设 E 是开集，则 E^c 是闭集；设 E 是闭集，则 E^c 是开集.

证 设 E 是开集，则 $E^0 = E$. 故 $\overline{E^c} = (E^0)^c = E^c$，所以 E^c 为闭集；设 E 是闭集，则 $\overline{E} = E$，故 $(E^c)^0 = (\overline{E})^c = E^c$，所以 E^c 为开集；

正是由于开集和闭集的这种对偶关系，在许多情况下，我们将闭集看作是由开集派生出来的一个概念. 也就是说如果定义了开集，闭集也就随之而确定.

定理 3 任意多个开集的并集是开集，有限多个开集的交集是开集.

证 设 $\{G_\alpha \mid \alpha \in \Gamma\}$ 是任一族开集，要证 $G = \bigcup_{\alpha \in \Gamma} G_\alpha$ 是开集. 设 $x \in G$，则存在 $\alpha \in \Gamma$，使得 $x \in G_\alpha$. 因为 G_α 是开集，所以存在 $O(x, \delta) \subseteq G_\alpha \subseteq G$，于是 $x \in G^0$，即 $G \subseteq G^0$，所以 G 是开集.

设 $G_1, G_2, \cdots, G_k$ 是开集，要证 $G = \bigcap_{i=1}^{k} G_i$ 是开集.

若 $G = \varnothing$，则 G 为开集；

若 $G \neq \varnothing$，任取 $x \in G$，则 $x \in G_i$ $(i = 1, 2, \cdots, k)$，从而有 $O(x, \delta_i) \subseteq G_i$ $(i = 1, 2, \cdots, k)$. 取 $\delta = \min_{1 \leqslant i \leqslant k}\{\delta_i\}$，有 $O(x, \delta) \subseteq \bigcap_{i=1}^{k} G_i = G$，故 $x \in G^0$，$G \subseteq G^0$，所以 G 是开集.

注意 无限多个开集的交集不一定是开集.

例如，$\bigcap_{i=1}^{\infty}\left(-\frac{1}{i}, \frac{1}{i}\right) = \{0\}$，每一个 $\left(-\frac{1}{i}, \frac{1}{i}\right)$ 都是 $\mathbf{E}^1$ 中的开集，而 $\{0\}$ 却不是 $\mathbf{E}^1$ 中的开集.

由定理 3 和 De Morgan 公式可知：

定理 4 任意多个闭集的交集是闭集，有限多个闭集的并集是闭集.

注意 无限多个闭集的并集不一定是闭集.

例如，$\left[\frac{1}{n}, 1 - \frac{1}{n}\right]$ $(n = 3, 4, \cdots)$ 全是 $\mathbf{E}^1$ 中的闭集，$\bigcup_{n=3}^{\infty}\left[\frac{1}{n}, 1 - \frac{1}{n}\right] = (0, 1)$，而 $(0, 1)$ 不是 $\mathbf{E}^1$ 中的闭集.

定义 3 设 E 是一个集合，$\mu = \{M_\alpha \mid \alpha \in \Gamma\}$ 是一族集合，若 $\bigcup_{\alpha \in \Gamma} M_\alpha \supseteq E$，则

说 μ 是 E 的一个覆盖.

当 Γ 是有限集或可数集时，分别说 μ 是 E 的有限覆盖或可数覆盖. 如果 μ 是 E 的覆盖，并且 μ 的子族 μ_1 也是 E 的覆盖，则说 μ_1 是 μ 关于 E 的子覆盖. 如果 $\mathbf{E}^n$ 中的开(闭)集族 μ 是 $\mathbf{E}^n$ 的子集 E 的覆盖，则说 μ 是 E 的开(闭)覆盖.

定理 5 (Heine-Borel **有限覆盖定理**)**$\mathbf{E}^n$ 中的有界闭集的每一开覆盖都有有限子覆盖.**

定理 5 的证明在这里略去，读者可借鉴数学分析中的有限覆盖定理的证明或参阅本篇末的参考文献.

定义 4 **设 $E \subseteq \mathbf{E}^n$，如果 $E \subseteq E'$，则说 E 是自密集，如果 $E' = E$，则说 E 是完备集.**

显然，完备集就是自密闭集，也就是没有孤立点的闭集.

例如，空集是完备集，$\mathbf{E}^1$ 中任一闭区间 $[a, b]$ 及全直线都是完备集.

表面看来，既然一个完备集一方面是闭集，而另一方面每一个点又都是聚点，似乎它就会铺满空间的一小块，但是这是一种错觉，下面的例子将会说明这一点.

例 (Cantor **集合**)将 $[0, 1]$ 进行三等分，删去中间的开区间 $\left(\frac{1}{3}, \frac{2}{3}\right)$，将剩下的两个闭区间分别进行三等分，各自删去中间的开区间，即 $\left(\frac{1}{9}, \frac{2}{9}\right)$ 和 $\left(\frac{7}{9}, \frac{8}{9}\right)$. 如此继续作下去，自然有些点是永远删不去的，例如 $\frac{1}{3}$ 和 $\frac{2}{3}$ 以及所有被删去的开区间的端点，所有这样的永远删不去的点作成的集合 E 就叫做 Cantor集合.

下面我们来证明 Cantor **集合** E **是一个完备集**：

首先证明 E 是一个闭集. 设 A 是所有被挖去的点作成的集合，则 A 是可数个开区间的并，所以是开集. 而 $E = [0, 1] - A = [0, 1] \cap A^c$，故 E 是闭集.

其次证明 E 是一个自密集. 注意到在进行第一次删去手续后所余下的两个闭区间的长度都是 $\frac{1}{3}$，进行第二次删去手续后剩下的四个闭区间的长度都是 $\frac{1}{3^2}$，在进行第 n 次删去手续后剩下来的 2^n 个闭区间的长度都是 $\frac{1}{3^n}$. 对于任意的 $x \in E$，任取 x 的邻域 $O(x, \delta)$，则对于取定的 $\delta > 0$，只要 n_0 取得充分大后就有 $\frac{1}{3^{n_0}} < \delta$，因为 x 是永远删不掉的点，则 x 应该属于删去 n_0 次后剩下的某一个闭区间 I_{n_0} 中，故 $I_{n_0} \subset O(x, \delta)$，于是 I_{n_0} 的两个端点也应该在 $O(x, \delta)$ 中，而它们也是 E 中的点，于是 $O(x, \delta)$ 中至少有一异于 x 的点属于 E，即是 $x \in E'$，所以

$E \subseteq E'$，即 E 是一个自密集.

故 Cantor 集合 E 是一个完备集.

注意 显然 Cantor 集合不能包含任意区间.

定义 5 设 $A \subseteq \mathbf{E}^n$，$E \subseteq \mathbf{E}^n$，若 E 中每一个点的任一邻域中都含有 A 的点，则说 A 在 E 中稠密. 若 $E = \mathbf{E}^n$，则说 A 是 $\mathbf{E}^n$ 中的稠密集. 若 A 不在 $\mathbf{E}^n$ 的任意一个非空开集中稠密，则说 A 是疏朗集.

疏朗集 E 的特征是 $\overline{E}$ 无内点.

Cantor **集是一个疏朗完备集.**

第四节 直线上的开集、闭集、完备集的构造

在本节中，我们将详细讨论直线上的开集，闭集，完备集的构造. 为此先引入构成区间的概念.

定义 1 设 G 是直线上的开集，如果开区间 $(\alpha, \beta) \subseteq G$，且 $\alpha, \beta \notin G$，则说 (α, β) 是 G 的一个构成区间.

例如，开集 $(0, 1) \cup (2, 3)$ 的构成区间是 $(0, 1)$ 及 $(2, 3)$.

定理 1 （开集的构造）直线上任意非空开集 G 可表成至多可数个互不相交的构成区间的并集，又，当非空开集表示成互不相交的开区间的并集时，这些区间必是构成区间.

证 设 G 是直线上的一个非空开集，分以下几步来论证：

(1) G 中任意一点必含在一个构成区间中. 对于任意 $x_0 \in G$，设 $F = [x_0, +\infty) \cap G^c$，于是 F 是一个非空有下界的闭集. 设 F 中最小的点是 β_0，则 $\beta_0 \geq x_0$. 因为 $x_0 \in G$，故 $x_0 \notin F$，因此只能有 $\beta_0 > x_0$. 而 $\beta_0 \notin G$，于是必有 $[x_0, \beta_0) \subseteq G$，如若不然，设 $y \in [x_0, \beta_0)$，但 $y \notin G$，因此有 $y \in F$，$y < \beta_0$，这与 β_0 是 F 中的最小点矛盾. 同样的可以证明存在点 α_0，满足 $\alpha_0 < x_0$，$\alpha_0 \notin G$，$(\alpha_0, x_0] \subseteq G$. 于是可得 G 的一个构成区间 (α_0, β_0) 包含着 x_0.

(2) G 的任意两个不同的构成区间必不相交. 不然的话，设 (α_1, β_1)，(α_2, β_2) 是 G 的两个不同的构成区间，但相交. 这时必有一个区间的端点在另一个区间内，例如 $\alpha_1 \in (\alpha_2, \beta_2)$，但 $(\alpha_2, \beta_2) \subseteq G$，这和 $\alpha_1 \notin G$ 矛盾，因此不同的构成区间不相交.

(3) G 的不同的构成区间至多可数个. 事实上，对于 G 的每一个构成区间，取其中的一个有理数与之对应，则可得 G 的构成区间全体与有理数的一个子集对等.

(4) 作 G 的所有构成区间的并集 $\bigcup(\alpha, \beta)$，由(1)它应是 G，由(2)和(3)，G 必定是有限个或可数个互不相交的构成区间的并集. 用 (α_ν, β_ν) $(\nu = 1, 2, \cdots, n$

或 $\nu = 1, 2, \cdots$) 记 G 的构成区间，那么 $G = \bigcup_{\nu} (\alpha_\nu, \beta_\nu)$.

(5) 设 $G = \bigcup_{\nu} (\alpha'_\nu, \beta'_\nu)$ 是一组互不相交的开区间的并集，现在只要证明每个 $(\alpha'_\nu, \beta'_\nu)$ 都是 G 的构成区间. 显然，$(\alpha'_\nu, \beta'_\nu) \subseteq G$. 若它不是构成区间，例如说 $\alpha'_\nu \in G$，那么必有 $\mu \neq \nu$，使得 $\alpha'_\nu \in (\alpha'_\mu, \beta'_\mu)$，因而 $(\alpha'_\mu, \beta'_\mu)$ 与 $(\alpha'_\nu, \beta'_\nu)$ 相交，这和假设矛盾. 所以 $\alpha'_\nu \notin G$. 同样 $\beta'_\nu \notin G$. 所以 $(\alpha'_\nu, \beta'_\nu)$ 是构成区间.

既然闭集的补集是开集，那么从开集的构成区间可以引入余区间的概念.

定义 2 **设 A 是直线上的闭集，A 的补集 A^c 的构成区间叫做 A 的余区间或邻接区间.**

定理 2 **(闭集的构造)直线上的闭集 F 或者是全直线，或者是从直线上挖掉有限个或可数个互不相交的开区间(即 F 的余区间)所得到的集.**

由孤立点的定义很容易知道，直线上点集 A 的孤立点必是包含在 A 的补集中的某两个开区间的公共端点，因此，闭集的孤立点一定是它的两个余区间的公共端点.

完备集是没有相邻接的余区间的闭集.

最后附带说明一下，当 $n > 1$ 时，$\mathbf{E}^n$ 中的开集一般不能表示成至多可数个互不相交的开区间(n 维)的并集，但总可表示成可数个互不相交的半开半闭(例如左开右闭)区间之并，不过这种表示法没有惟一性(在 $\mathbf{E}^1$ 中一个开集只能用一种方式表示成构成区间之并).

习　题　1-2

1. 证明：$x_0 \in E'$ 的充分必要条件是在任意含有 x_0 的邻域 $O(x, \delta)$ 中，恒有异于 x_0 而属于 E 的点存在. 而 $x_0 \in E^\circ$ 的充分必要条件是有含有 x_0 的邻域 $O(x, \delta)$ 存在，使得 $O(x, \delta) \subseteq E$.

2. 设 E 是$[0, 1]$中的全部有理点所成之集，在 $\mathbf{E}^1$ 与 $\mathbf{E}^2$ 中分别看 E 时，求 E', E°, $\overline{E}$.

3. 设 $\mathbf{E}^n = \mathbf{E}^2$ 是普通的 xOy 平面，

(1) $E_1 = \{(x, y) \mid x^2 + y^2 < 1\}$，求 E_1', E_1°, $\overline{E_1}$；

(2) $E_2 = \left\{(x, y) \mid y = \begin{cases} \sin \dfrac{1}{x}, & \text{当 } x \neq 0 \text{ 时}, \\ 0, & \text{当 } x = 0 \text{ 时}, \end{cases} x \in \mathbf{R}\right\}$，求 E_2', E_2°.

4. 每个闭集均是可数个开集的交集，每个开集可以表示成可数个闭集的并集.

5. 证明：$f(x)$为$[a, b]$上连续函数的充分必要条件是对任意实数 c，集 $E = \{x \mid f(x) \geqslant c\}$ 和 $E_1 = \{x \mid f(x) \leqslant c\}$ 都是闭集.

6. 任何闭区间不可能表为可数个疏朗集的并集.

第三章　Lebesgue 测度

实变函数论的中心问题是建立一种新的积分——Lebesgue 积分的理论.从本章起我们进入新积分理论的第一步,即研究与新积分相关的点集的"测度".

测度概念在 $\mathbf{E}^1$ 中是长度概念的推广,在 $\mathbf{E}^2$ 中是面积概念的推广,而且可以超出 $\mathbf{E}^n$ 在抽象空间中定义测度.测度论不只是近代积分论的基础,在其他数学分支里也有广泛的应用.本章主要介绍 Lebesgue 测度(简称 L 测度).

第一节　外测度　内测度

定义 1　设 E 是 $\mathbf{E}^n$ 中一点集,对于每一列覆盖 E 的开区间 $\bigcup_{i=1}^{\infty} I_i \supseteq E$,记

$$m^* E = \inf_{E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} |I_i|,$$

则 $m^* E$ 叫做集合 E 的Lebesgue 外测度,简称L 外测度或外测度.

外测度具有以下基本性质:

定理 1　(1) $m^* E \geqslant 0,\ m^* \varnothing = 0$　　　　(非负性);

(2) 设 $A \subseteq B$,则 $m^* A \leqslant m^* B$　　　　(单调性);

(3) $m^* (\bigcup_{i=1}^{\infty} A_i) \leqslant \sum_{i=1}^{\infty} m^* A_i$　　　　(次可数可加性);

(4) 若 $d(A, B) > 0$,则 $m^* (A \cup B) = m^* A + m^* B$.

证　(1) 显然成立.

(2) 因为 $A \subseteq B$,所以任一列覆盖 B 的开区间 $\{I_i\}$ 必然是覆盖 A 的,因而有 $m^* A \leqslant \sum_{i=1}^{\infty} |I_i|$.对所有能覆盖 B 的开区间列取体积和的下确界即得

$$m^* A \leqslant \inf_{\bigcup_{i=1}^{\infty} I_i \supseteq B} \sum_{i=1}^{\infty} |I_i| = m^* B.$$

(3) $\forall \varepsilon > 0$,由外测度的定义,对每个 n 都应有一列开区间 $I_{n1}, I_{n2}, \cdots, I_{nm}, \cdots$ 使得 $A_n \subseteq \bigcup_{m=1}^{\infty} I_{nm}$,且

$$\sum_{m=1}^{\infty} |I_{nm}| < m^* A_n + \frac{\varepsilon}{2^n},$$

从而

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n,m=1}^{\infty} I_{nm}$$

且
$$\sum_{n,m=1}^{\infty} |I_{nm}| = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |I_{nm}| < \sum_{n=1}^{\infty}\left(m^* A_n + \frac{\varepsilon}{2^n}\right)$$

$$= \sum_{n=1}^{\infty} m^* A_n + \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \sum_{n=1}^{\infty} m^* A_n + \varepsilon.$$

可见
$$m^*(\bigcup_{n=1}^{\infty} A_n) < \sum_{n,m=1}^{\infty} |I_{nm}| < \sum_{n=1}^{\infty} m^* A_n + \varepsilon.$$

由 ε 的任意性，有 $m^*(\bigcup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} m^* A_n$.

(4) 设 $\gamma = d(A, B) > 0$. 对于任意的正数 ε，由外测度的定义，存在一列开区间覆盖了 A 和 B，即 $(A \cup B) \subseteq \bigcup_{i=1}^{\infty} I_i$，且 $\sum_{i=1}^{\infty} |I_i| < m^*(A \cup B) + \varepsilon$. 由于每个开区间 I_i 都可以分解成至多可数个互不相交的直径小于 γ 的左开右闭区间 I_i^j 之并，并且 $|I_i| = \sum_j |I_i^j|$. 因此，不妨设 $\{I_i\}$ 中的每个区间的直径都小于 γ，这样的 I_i 显然不能同时含有 A 和 B 中的点. 将这列开区间中含有 A 中点的记为 $\{I_{i_k}\}$，不含有 A 中点的记为 $\{I_{i_m}\}$，则 $A \subseteq \bigcup_k I_{i_k}$，$B \subseteq \bigcup_m I_{i_m}$，于是

$$m^*(A \cup B) + \varepsilon > \sum_{i=1}^{\infty} |I_i| = \sum_k |I_{i_k}| + \sum_m |I_{i_m}| \geqslant m^* A + m^* B.$$

由 $\varepsilon > 0$ 的任意性知

$$m^*(A \cup B) \geqslant m^* A + m^* B.$$

另一方面，由(3)有

$$m^*(A \cup B) \leqslant m^* A + m^* B,$$

故

$$m^*(A \cup B) = m^* A + m^* B.$$

例 1 设 $A = \{x_0\}$ 是 $\mathbf{E}^n$ 中的独点集，其中 $x_0 = \{x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}\}$，则 $m^* A = 0$.

证 对于任意的正数 ε，令

$$I=\left\{(x_1, x_2, \cdots, x_n)\left| x_i^{(0)}-\frac{\varepsilon^{\frac{1}{n}}}{2}<x_i<x_i^{(0)}+\frac{\varepsilon^{\frac{1}{n}}}{2}, i=1, 2, \cdots, n\right.\right\},$$

则 $A\subset I$，且 $|I|=\varepsilon$. 从而有 $0\leqslant m^*A\leqslant\varepsilon$. 由 $\varepsilon>0$ 的任意性知 $m^*A=0$.

例 2 对任意区间 I，有 $m^*I=|I|$.

证 若 I 是闭区间，则对于任给的正数 ε，存在开区间 I_0，使得 $I\subset I_0$ 且 $|I_0|<|I|+\varepsilon$，由外测度的定义可知 $m^*I<|I|+\varepsilon$，再由 ε 的任意性，有 $m^*I\leqslant|I|$.

而对于上述任给的正数 ε，同时也存在一列开区间 $\{I_i\}$，使得 $I\subseteq\bigcup\limits_{i=1}^{\infty}I_i$ 且 $\sum\limits_{i=1}^{\infty}|I_i|<m^*I+\varepsilon$. 由有限覆盖定理知在 $\{I_i\}$ 中存在有限多个区间，不妨设为 $I_1, I_2, \cdots, I_n$，使得 $I\subseteq\bigcup\limits_{i=1}^{n}I_i$. 由于 $I=\bigcup\limits_{i=1}^{n}(I\cap I_i)$，由初等几何的知识可知 $|I|\leqslant\sum\limits_{i=1}^{n}|I\cap I_i|$，故 $|I|\leqslant\sum\limits_{i=1}^{n}|I\cap I_i|\leqslant\sum\limits_{i=1}^{n}|I_i|<m^*I+\varepsilon$，由 ε 的任意性即得 $m^*I\geqslant|I|$.

综上可知 $m^*I=|I|$.

若 I 是其他任意的区间，则可作闭区间 I_1, I_2，使得

$$I_1\subseteq I\subseteq I_2 \text{ 且 } |I_2|-\varepsilon<|I|<|I_1|+\varepsilon,$$

于是

$$|I|-\varepsilon<|I_1|=m^*I_1\leqslant m^*I\leqslant m^*I_2=|I_2|<|I|+\varepsilon,$$

由 ε 的任意性即得 $m^*I=|I|$.

定理 2 设 E 是 $\mathbf{E}^n$ 中任一点集，则

$$m^*E=\inf\{m^*G\mid G \text{ 是包含 } E \text{ 的开集}\}.$$

证 分以下两步进行证明：

(1) 由外测度的单调性知，对任意包含 E 的开集 G，有 $m^*G\geqslant m^*E$.

(2) 若 $m^*E=+\infty$，则由(1)可推知定理成立. 因此，不妨设 $m^*E<+\infty$，于是，对于任意的正数 ε，由外测度的定义知，存在一列开区间 $I_i\ (i=1, 2, \cdots)$，使得 $E\subseteq\bigcup\limits_{i=1}^{\infty}I_i$，且 $\sum\limits_{i=1}^{\infty}|I_i|<m^*E+\varepsilon$.

令 $G=\bigcup\limits_{i=1}^{\infty}I_i$，则 G 是开集，$G\supseteq E$，且

$$m^*G=m^*(\bigcup_{i=1}^{\infty}I_i)\leqslant\sum_{i=1}^{\infty}m^*I_i=\sum_{i=1}^{\infty}|I_i|<m^*E+\varepsilon.$$

由(1)(2)两步即知，定理成立.

定义 2 设 E 为 $\mathbf{E}^n$ 中有界集，I 为任一包含 E 的开区间，则 $|I|-m^*(I-E)$ 叫做 E 的内测度，记为 m_*E.

注意 由于 $m^*I=|I|$ 及外测度的单调性和次可加性，总有 $m_*E\geqslant 0$ 及 $m_*E\leqslant m^*E$.

第二节 可 测 集

一、可测集的定义

定义 1 设 E 为 $\mathbf{E}^n$ 中有界集，如果 $m^*E=m_*E$，则说 E 是Lebesgue 可测的，简称L 可测. 又设 E 是 $\mathbf{E}^n$ 中的无界集，如果对任意开区间 I，有界集 $E\cap I$ 都是 L 可测的，则说 E 是 L 可测的.

如果 E 是 $\mathbf{E}^n$ 中 L 可测集，m^*E叫做 E 的 L 测度，简记为mE.

以上是 Lebesgue 下的定义，在定义中区分了有界集与无界集，而且同时出现了两种测度，用起来不方便，下面一个形式比较简洁的等价定义是由 Carathéodory 给出的.

定义 2 设 E 为 $\mathbf{E}^n$ 中的点集，如果对任一点集 T 都有

$$m^*T=m^*(T\cap E)+m^*(T\cap E^c),$$

则说 E 是Lebesgue 可测的，简称为L 可测. 这时，E 的 L 外测度m^*E就叫做 E 的L 测度，记为mE.

这个定义没有上述定义的缺点，以后就用它作为可测集的定义.

二、可测集的性质

定理 1 集 E 可测的充分必要条件是对于任意的 $A\subseteq E$, $B\subseteq E^c$，恒有

$$m^*(A\cup B)=m^*A+m^*B.$$

证 必要性：取 $T=A\cup B$，则

$$T\cap E=(A\cup B)\cap E=A,$$

$$T\cap E^c=(A\cup B)\cap E^c=B,$$

于是

$$m^*(A\cup B)=m^*T=m^*(T\cap E)+m^*(T\cap E^c)=m^*A+m^*B.$$

充分性：对任意点集 T，令 $A=T\cap E$, $B=T\cap E^c$，则

$$A \subseteq E,\ B \subseteq E^c,$$
$$A \cup B = (T \cap E) \cup (T \cap E^c) = T \cap (E \cup E^c) = T.$$

由于 $m^*(A \cup B) = m^*A + m^*B$,所以 $m^*T = m^*(T \cap E) + m^*(T \cap E^c)$,这说明 E 可测.

定理 2　集 E 可测的充分必要条件是 E^c 可测.

证　由可测的定义及 $(E^c)^c = E$,立即可得此定理.

定理 3　若集 E_1, E_2 都可测,则 $E_1 \cap E_2$ 也可测.

证　因集 E_1, E_2 可测,则对任意点集 T 有

$$m^*T = m^*(T \cap E_1) + m^*(T \cap E_1^c),$$
$$\begin{aligned} m^*(T \cap E_1) &= m^*[(T \cap E_1) \cap E_2] + m^*[(T \cap E_1) \cap E_2^c] \\ &= m^*[T \cap (E_1 \cap E_2)] + m^*[T \cap E_1 \cap E_2{}^c], \end{aligned}$$
$$\begin{aligned} m^*[T \cap (E_1 \cap E_2)^c] &= m^*[(T \cap (E_1 \cap E_2)^c) \cap E_1] \\ &\quad + m^*[(T \cap (E_1 \cap E_2)^c) \cap E_1^c] \\ &= m^*[T \cap E_1 \cap E_2^c] + m^*(T \cap E_1^c), \end{aligned}$$

因此

$$\begin{aligned} m^*T &= m^*(T \cap E_1) + m^*(T \cap E_1^c) \\ &= m^*[T \cap (E_1 \cap E_2)] + m^*[T \cap E_1 \cap E_2^c] \\ &\quad + m^*(T \cap E_1^c) \\ &= m^*[T \cap (E_1 \cap E_2)] + m^*[T \cap (E_1 \cap E_2)^c], \end{aligned}$$

故 $E_1 \cap E_2$ 可测.

推论 1　若集 E_i $(i = 1, 2, \cdots, k)$ 都可测,则 $\bigcap\limits_{i=1}^{k} E_i$ 也可测.

定理 4　若集 E_1, E_2 都可测,则 $E_1 - E_2$ 也可测.

证　因 $E_1 - E_2 = E_1 \cap E_2^c$,由定理 2 及定理 3 知,$E_1 - E_2$ 可测.

定理 5　若集 E_1, E_2 都可测,则 $E_1 \cup E_2$ 也可测,且若 $E_1 \cap E_2 = \varnothing$,则对任意点集 T,恒有

$$m^*[T \cap (E_1 \cup E_2)] = m^*(T \cap E_1) + m^*(T \cap E_2).$$

证　因 $E_1 \cup E_2 = (E_1^c \cap E_2^c)^c$,故 $E_1 \cup E_2$ 可测.

若 $E_1 \cap E_2 = \varnothing$,则 $T \cap E_1 \subseteq E_1$, $T \cap E_2 \subseteq E_2 \subseteq E_1^c$,由定理 1 知,

$$\begin{aligned} m^*[T \cap (E_1 \cup E_2)] &= m^*[(T \cap E_1) \cup (T \cap E_2)] \\ &= m^*(T \cap E_1) + m^*(T \cap E_2). \end{aligned}$$

推论 2 若集 E_i $(i=1,2,\cdots,k)$ 都可测，则 $\bigcup\limits_{i=1}^{k}E_i$ 也可测，且若 $E_i\cap E_j=\varnothing(i\neq j)$，则对于任意点集 T，恒有

$$m^*[T\cap(\bigcup_{i=1}^{k}E_i)]=\sum_{i=1}^{k}m^*(T\cap E_i).$$

注意 在 $m^*[T\cap(\bigcup\limits_{i=1}^{k}E_i)]=\sum\limits_{i=1}^{k}m^*(T\cap E_i)$ 中取 $T=\bigcup\limits_{i=1}^{k}E_i$，就得到测度的有限可加性

$$m(\bigcup_{i=1}^{k}E_i)=\sum_{i=1}^{k}E_i.$$

推论 3 设集 E_1，E_2 都可测，且 $E_1\subseteq E_2$，$mE_1<+\infty$，则

$$m(E_2-E_1)=mE_2-mE_1.$$

证 因 $mE_2=m[(E_2-E_1)\cup E_1]=m(E_2-E_1)+mE_1$，
故 $m(E_2-E_1)=mE_2-mE_1$.

定理 6 （测度的可数可加性）设集 E_i $(i=1,2,\cdots)$ 都可测，$E_i\cap E_j=\varnothing(i\neq j)$，则 $\bigcup\limits_{i=1}^{\infty}E_i$ 也可测，且 $m(\bigcup\limits_{i=1}^{\infty}E_i)=\sum\limits_{i=1}^{\infty}mE_i$.

证 一方面，由推论 2 知，对任意正整数 k，$\bigcup\limits_{i=1}^{k}E_i$ 总是可测的，因此对任意点集 T，恒有

$$\begin{aligned}m^*T&=m^*[T\cap(\bigcup_{i=1}^{k}E_i)]+m^*[T\cap(\bigcup_{i=1}^{k}E_i)^c]\\&=\sum_{i=1}^{k}m^*(T\cap E_i)+m^*[T\cap(\bigcup_{i=1}^{k}E_i)^c]\\&\geqslant\sum_{i=1}^{k}m^*(T\cap E_i)+m^*[T\cap(\bigcup_{i=1}^{\infty}E_i)^c],\end{aligned}$$

令 $k\to\infty$，即得

$$\begin{aligned}m^*T&\geqslant\sum_{i=1}^{\infty}m^*(T\cap E_i)+m^*[T\cap(\bigcup_{i=1}^{\infty}E_i)^c]\\&\geqslant m^*[T\cap(\bigcup_{i=1}^{\infty}E_i)]+m^*[T\cap(\bigcup_{i=1}^{\infty}E_i)^c].\end{aligned}$$

另一方面，因为

$$T=[T\cap(\bigcup_{i=1}^{\infty}E_i)]\cup[T\cap(\bigcup_{i=1}^{\infty}E_i)^c],$$

故

$$m^* T \leqslant m^*[T \cap (\bigcup_{i=1}^{\infty} E_i)] + m^*[T \cap (\bigcup_{i=1}^{\infty} E_i)^c].$$

由以上两个方面知

$$m^* T = m^*[T \cap (\bigcup_{i=1}^{\infty} E_i)] + m^*[T \cap (\bigcup_{i=1}^{\infty} E_i)^c],$$

这说明 $\bigcup_{i=1}^{\infty} E_i$ 可测.

取 $T = \bigcup_{i=1}^{\infty} E_i$，则 $T \cap (\bigcup_{i=1}^{\infty} E_i)^c = \varnothing$，$T \cap E_i = E_i$，$i = 1, 2, \cdots$，

由 $m^* T \geqslant \sum_{i=1}^{\infty} m^*(T \cap E_i) + m^*[T \cap (\bigcup_{i=1}^{\infty} E_i)^c]$ 知

$$m(\bigcup_{i=1}^{\infty} E_i) \geqslant \sum_{i=1}^{\infty} mE_i,$$

再由第一节定理 1 的(3)知

$$m(\bigcup_{i=1}^{\infty} E_i) \leqslant \sum_{i=1}^{\infty} mE_i,$$

故

$$m(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} mE_i.$$

推论 4 **设集 $E_i (i = 1, 2, \cdots)$ 都可测，则 $\bigcup_{i=1}^{\infty} E_i$ 也可测.**

证 令 $E_1^* = E_1$，$E_i^* = E_i - (\bigcup_{j=1}^{i-1} E_j)$ $(i = 2, 3, \cdots)$，则 E_i^* $(i = 1, 2, 3, \cdots)$ 都可测，且 $E_i^* \cap E_j^* = \varnothing (i \neq j)$，又 $\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} E_i^*$，故 $\bigcup_{i=1}^{\infty} E_i$ 可测.

定理 7 **设集 $E_i (i = 1, 2, \cdots)$ 都可测，则 $\bigcap_{i=1}^{\infty} E_i$ 也可测.**

证 因 $\bigcap_{i=1}^{\infty} E_i = (\bigcup_{i=1}^{\infty} E_i^c)^c$，故由推论 4 与定理 2，$\bigcap_{i=1}^{\infty} E_i$ 可测.

定理 8 **设集 $\{E_k\} (k = 1, 2, \cdots)$ 是单调增加的可测集列，$E = \bigcup_{k=1}^{\infty} E_k = \lim_{k \to \infty} E_k$，则**

$$\lim_{k \to \infty} mE_k = m(\lim_{k \to \infty} E_k) = mE \quad \text{（下连续性）}.$$

证 由推论 4，$E = \bigcup_{i=1}^{\infty} E_i$ 可测. 设 $E_0 = \varnothing$，$A_i = E_i - E_{i-1} (i = 1, 2, \cdots)$，则 A_i $(i = 1, 2, \cdots)$ 可测，$A_i \cap A_j = \varnothing (i \neq j)$，且 $E_k = \bigcup_{i=1}^{k} A_i$，$E = \bigcup_{i=1}^{\infty} A_i$.

由定理 6 及推论 2 的注意，

$$mE = \sum_{i=1}^{\infty} mA_i = \lim_{k\to\infty} \sum_{i=1}^{k} mA_i = \lim_{k\to\infty} m(\bigcup_{i=1}^{k} A_i) = \lim_{k\to\infty} mE_k.$$

定理 9 设 $\{E_k\}(k = 1, 2, \cdots)$ 是单调减少的可测集列，$E = \bigcap_{k=1}^{\infty} E_k = \lim_{k\to\infty} E_k$，且 $mE_1 < +\infty$，则

$$\lim_{k\to\infty} mE_k = m(\lim_{k\to\infty} E_k) = mE \quad \text{（上连续性）}.$$

证 由定理 7，$E = \bigcap_{i=1}^{\infty} E_i$ 可测. 由于 $\{E_1 - E_k\}$ 是单调增加的可测集列，所以

$$\lim_{k\to\infty}(E_1 - E_k) = \bigcup_{i=1}^{\infty} (E_1 - E_i) = E_1 - (\bigcap_{i=1}^{\infty} E_i) = E_1 - E,$$

又由于 $E \subseteq E_k \subseteq E_1 (k \in \mathbf{Z}^+)$，从而

$$mE \leqslant mE_k \leqslant mE_1 < +\infty (k \in \mathbf{Z}^+),$$

由推论 3 和定理 8，

$$\begin{aligned} mE_1 - \lim_{k\to\infty} mE_k &= \lim_{k\to\infty}(mE_1 - mE_k) \\ &= \lim_{k\to\infty} m(E_1 - E_k) = m(E_1 - E) \\ &= mE_1 - mE, \end{aligned}$$

故

$$\lim_{k\to\infty} mE_k = mE.$$

注意 定理 9 中 $mE_1 < +\infty$ 的条件可以改为存在某个正整数 k_0，使得 $mE_{k_0} < +\infty$，但此条件不能去掉.

三、常见的可测集合

定理 10 (1) 凡外测度为零之集皆可测，叫做零测度集；

(2) 零测度集之任意子集是零测度集；

(3) 至多可数个零测度集的并集是零测度集.

定理 11 区间 I 都是可测集，且 $mI = |I|$.

以上两个定理的证明留作习题.

定理 12 凡开集、闭集皆可测.

证 这是因为任意非空开集皆可表示为至多可数个互不相交的左开右闭区

间之并，而区间是可测的；开集可测，则闭集作为开集的补集自然可测.

定义 3 设集合 G 可表示为一列开集 $\{G_i\}$ 的交集，则 G 叫做 G_δ 型集；设集合 F 可表示为一列闭集 $\{F_i\}$ 的并集，则 F 叫做 F_σ 型集.

定义 4 凡属可以从开集出发，用取补集、取至多可数个集合的并集或交集等手续不超过可数多次而得到的集合叫做 Borel 集.

显然 G_δ 型集与 F_σ 型集都是 Borel 集，且有：

定理 13 凡 Borel 集都可测.

定理 14 设 E 是任一可测集，则一定存在 G_δ 型集 G，使得 $G \supseteq E$，且 $m(G-E)=0$，同时也一定存在 F_σ 型集 F，使得 $F \subseteq E$，且 $m(E-F)=0$.

上面的定理告诉我们，只要有了全部 G_δ 型集 G（或全部 F_σ 型集 F）和全部 L 零测度集 N，那么一切 L 可测集都可以获得，它们一律可以表示成 $E=G-N$ 或 $E=F\cup N$ 的形式.

最后介绍一个以后常用的概念.

定义 5 设 Π 是一个与集合 E 的点 x 有关的命题，如果

$$E-E[\Pi\text{ 成立}] = \text{零测度集},$$

则说命题 Π 在 E 上几乎处处成立，记为 Π a. e. 于 E.

习　题　1-3

1. 证明本章第二节定理 10 与定理 11.

2. 证明：若 E 有界，则 $m^*E<+\infty$.

3. 证明：至多可数集的外测度为零.

4. E 为直线上的有界集，$m^*E>0$，则对于任一小于 m^*E 的正数 p，存在 E 的子集 E_1，使得 $m^*E_1=p$.

5. 证明：Cantor 集为零测度集.

6. 设 $\{E_k\}$ 是 $[0,1]$ 中的可测集列，$mE_k=1\ (k=1,2,\cdots)$，证明：

$$m\left(\bigcap_{k=1}^{\infty}E_k\right)=m\left(\bigcup_{k=1}^{\infty}E_k\right)=1.$$

7. 设 $E_1\subseteq E_2$，$E_i\subseteq \mathbf{E}^n\ (i=1,2)$，$E_1$ 可测，且 $mE_1=m^*E_2$，则 E_2 为可测集.

8. 若 E 可测，则对于任意正数 ε，恒有开集 G 及闭集 F，使得 $F\subseteq E\subseteq G$，且 $m(G-E)<\varepsilon$，$m(E-F)<\varepsilon$.

第四章 可测函数

为了建立 Lebesgue 积分，我们还必须讨论一类重要的函数——Lebesgue 可测函数，它也是从测度观点来研究函数时所必然要考虑的一类函数. 我们将看到，在可测函数类中进行运算，如代数运算、取极限运算等都是相当方便的，且所得结果仍然是可测函数.

为了今后讨论问题的需要，我们先给出广义实数集及广义实值函数等概念，并对一些常用的与函数有关的集合的记号作出规定.

集合 $\hat{\mathbf{R}} = (-\infty, +\infty) \cup \{-\infty\} \cup \{+\infty\}$ 叫做广义实数集，并规定 $+\infty$，$-\infty$与实数 $x(-\infty < x < +\infty)$ 的代数运算如下：

(1) 加法 $x + (\pm\infty) = (\pm\infty) + x = (\pm\infty) + (\pm\infty) = \pm\infty$；

(2) 减法 $x - (\mp\infty) = (\pm\infty) - x = \pm\infty$；

(3) 乘法 $(\pm\infty)(\pm\infty) = +\infty$，$(\pm\infty)(\mp\infty) = -\infty$，

$$x(\pm\infty) = (\pm\infty)x = \begin{cases} \pm\infty, & \text{当 } x > 0 \text{ 时；} \\ 0, & \text{当 } x = 0 \text{ 时；} \\ \mp\infty, & \text{当 } x < 0 \text{ 时；} \end{cases}$$

(4) 除法 $\dfrac{x}{\pm\infty} = 0$，$\dfrac{\pm\infty}{x} = \begin{cases} \pm\infty, & \text{当 } x > 0 \text{ 时；} \\ \mp\infty, & \text{当 } x < 0 \text{ 时.} \end{cases}$

注意 $(\pm\infty) + (\mp\infty)$，$(\pm\infty) - (\pm\infty)$，$\dfrac{\pm\infty}{0}$，$\dfrac{x}{0}$，$\dfrac{\pm\infty}{\pm\infty}$，$\dfrac{\mp\infty}{\pm\infty}$ 都是没有意义的.

E 是任意非空集合，把映射 $f: E \to \hat{\mathbf{R}}$ 叫做 E 上的广义实值函数，而把映射 $g: E \to \mathbf{R}$ 叫做 E 上的实值函数或有限函数.

如果 $f_n(x)$ $(n = 1, 2, \cdots)$，$f(x)$，$g(x)$ 都是定义在集合 E 上的广义实值函数，a，b 是有限实数，我们记

$$E[f > a] = \{x \mid x \in E,\ f(x) > a\},$$

$$E[a < f \leqslant b] = \{x \mid x \in E,\ a < f(x) \leqslant b\},$$

$$E[f \neq g] = \{x \mid x \in E,\ f(x) \neq g(x)\},$$

$$E[f_n \to f] = \{x \mid x \in E,\ \lim_{n\to\infty} f_n(x) = f(x)\},$$

等等，这些集合都是 E 的子集. 今后我们利用集合来分析函数的性质时常要用

到这些类型的集合.

第一节　可测函数及其性质

一、可测函数

定义 1　设 $f(x)$ 是定义在可测集 $E\subseteq\mathbf{E}^n$ 上的广义实值函数，$\forall a\in\mathbf{R}$，$E[f>a]$ 都是可测集，则说 $f(x)$ 是定义在 E 上的可测函数.

定理 1　设 $f(x)$ 是定义在可测集 $E\subseteq\mathbf{E}^n$ 上的广义实值函数，下列任一条件都是 $f(x)$ 在 E 上可测的充分必要条件：

(1) $\forall a\in\mathbf{R}$，$E[f\geqslant a]$ 都可测；

(2) $\forall a\in\mathbf{R}$，$E[f<a]$ 都可测；

(3) $\forall a\in\mathbf{R}$，$E[f\leqslant a]$ 都可测；

(4) $\forall a,b\in\mathbf{R}$，且 $a<b$，$E[a\leqslant f<b]$ 都可测(但充分性要假定 $|f(x)|<+\infty$).

证　$E[f\geqslant a]$ 与 $E[f<a]$ 对于 E 是互补的，同样 $E[f\leqslant a]$ 与 $E[f>a]$ 对于 E 也是互补的，故只需证明(1)与(4).

注意到

$$E[f\geqslant a]=\bigcap_{n=1}^{\infty}E\left[f>a-\frac{1}{n}\right],$$

$$E[f>a]=\bigcup_{n=1}^{\infty}E\left[f\geqslant a+\frac{1}{n}\right]$$

与

$$E[a\leqslant f<b]=E[f\geqslant a]-E[f\geqslant b],$$

$$E[f\geqslant a]=\bigcup_{n=1}^{\infty}E[a\leqslant f<a+n]$$

即可得证.

推论 1　设 $f(x)$ 是定义在 $E\subseteq\mathbf{E}^n$ 上的可测函数，则 $E[f=a]$ 是可测集，不论 a 是有限实数还是 $\pm\infty$.

证　只须注意　$E[f=a]=E[f\geqslant a]-E[f>a]$，

$$E[f=+\infty]=\bigcap_{n=1}^{\infty}E[f>n],$$

$$E[f=-\infty]=\bigcap_{n=1}^{\infty}E[f<-n].$$

例 1　区间 $[a,b]$ 上的单调函数是可测函数.

证　不妨设 $f(x)$ 是 $E=[a,b]$ 上的单调增加函数.

$\forall c\in\mathbf{R}$，令

$$x_c=\sup_{x\in E}\{x\mid f(x)\leqslant c\}.$$

由于 $f(x)$ 是单调增加函数，故有

$$E[f>c]=(x_c,+\infty)\cap E,$$

所以 $E[f>c]$ 可测，说明了 $f(x)$ 是可测函数.

二、可测函数的性质

引理　**设 $f(x)$，$g(x)$ 都是定义在集合 $E\subseteq\mathbf{E}^n$ 上的可测函数，则 $E[f>g]$ 与 $E[f\geqslant g]$ 都是可测集.**

证　因为 $E[f\geqslant g]=E-E[f<g]$，故只需证明 $E[f>g]$ 可测.

设 $x_0\in E[f>g]$，于是有 $f(x_0)>g(x_0)$，则必存在有理数 r，使得 $f(x_0)>r>g(x_0)$，即是 $x_0\in E[f>r]\cap E[g<r]$，反之亦然.

因此，设有理数全体为 $r_1,r_2,\cdots$，则

$$E[f>g]=\bigcup_{n=1}^{\infty}(E[f>r_n]\cap E[g<r_n]),$$

由定理1可知等式右边是可测集.

定理2　**设 $f(x)$，$g(x)$ 在 $E\subseteq\mathbf{E}^n$ 上可测，则定义在 E 上的函数**

$$f(x)+g(x);\ |f(x)|;\ \frac{1}{f(x)};\ f(x)g(x)$$

都在 E 上可测.

证　$\forall a\in\mathbf{R}$，因为

$$E[f+c>a]=E[f>a-c],$$

$$E[cf>a]=\begin{cases}E\left[f>\dfrac{a}{c}\right], & \text{当 } c>0 \text{ 时;}\\ E\left[f<\dfrac{a}{c}\right], & \text{当 } c<0 \text{ 时}\end{cases}\quad(c\text{ 是非零有限常数}),$$

故 $f(x)+c$，$cf(x)$ 是可测的（显然当 c 等于零时这两个函数也是可测的）.

再注意到

$$E[f+g>a]=E[f>-g+a],$$

$$E[|f|>a]=\begin{cases}E[f>a]\cup E[f<-a], & \text{当 } a\geqslant 0 \text{ 时;}\\ E, & \text{当 } a<0 \text{ 时,}\end{cases}$$

$$E\left[\frac{1}{f}>a\right]=\begin{cases}E[f>0]\cap E\left[f<\frac{1}{a}\right], & \text{当 } a>0 \text{ 时；}\\ E[f>0]-E[f=+\infty], & \text{当 } a=0 \text{ 时；}\\ E[f>0]\cup E\left[f<\frac{1}{a}\right], & \text{当 } a<0 \text{ 时，}\end{cases}$$

$$E[f\cdot g>a]=\begin{cases}\{(E[f>0]\cap E[g>0])\cup(E[f<0] \\ \cap E[g<0])\}\cap E\left[|f|>\frac{a}{|g|}\right], & \text{当 } a\geqslant 0 \text{ 时；}\\ E-E[f\cdot g\leqslant a]=E-\Big\{(E[f>0]\cap E[g<0]) \\ \cup(E[f<0]\cap E[g>0])\cap E\left[|f|\geqslant\frac{-a}{|g|}\right]\Big\}, & \text{当 } a<0 \text{ 时，}\end{cases}$$

可知结论成立.

定理 3 设 $\{f_n(x)\}$ 是 $E\subseteq \mathbf{E}^n$ 上一列(或有限个)可测函数，则 $\mu(x)=\inf\limits_n f_n(x)$ 与 $\lambda(x)=\sup\limits_n f_n(x)$ 都在 E 上可测.

证 由 $E[\mu\geqslant a]=\bigcap\limits_n E[f_n\geqslant a]$ 与 $E[\lambda\leqslant a]=\bigcap\limits_n E[f_n\leqslant a]$ 可知结论成立.

定理 4 设 $\{f_n(x)\}$ 是 $E\subseteq \mathbf{E}^n$ 上一列可测函数，则 $F(x)=\varliminf\limits_{n\to\infty} f_n(x)$ 与 $G(x)=\varlimsup\limits_{n\to\infty} f_n(x)$ 也在 E 上可测.

证 由 $\varliminf\limits_{n\to\infty} f_n(x)=\sup\limits_n(\inf\limits_{m\geqslant n} f_m(x))$ 和 $\varlimsup\limits_{n\to\infty} f_n(x)=\inf\limits_n(\sup\limits_{m\geqslant n} f_m(x))$，再重复运用定理 3 可知结论成立.

推论 2 设 $\{f_n(x)\}$ 是 $E\subseteq \mathbf{E}^n$ 上一列可测函数，$\lim\limits_{n\to\infty} f_n(x)=f(x)$，则 $f(x)$ 在 E 上可测.

定理 5 (1) 设 $f(x)$ 是定义在可测集 $E\subseteq \mathbf{E}^n$ 上的可测函数，E_1 是 E 的可测子集，则 $f(x)$ 看作定义在 E_1 上的函数时是 E_1 上的可测函数；

(2) 设 $f(x)$ 是定义在可测集 $E_i(i=1,2,\cdots,s)$ 上的可测函数，则 $f(x)$ 在 $\bigcup\limits_{i=1}^{s} E_i$ 上可测.

证 由 $E_1[f>a]=E_1\cap E[f>a]$ 与 $(\bigcup\limits_{i=1}^{s} E_i)[f>a]=\bigcup\limits_{i=1}^{s} E_i[f>a]$ 可知结论成立.

下面介绍一类重要的可测函数.

定义 2 设 $f(x)$ 是定义在可测集 $E\subseteq \mathbf{E}^n$ 上的实值函数，若有 $E=\bigcup\limits_{i=1}^{s} E_i$，$E_i\cap E_j=\varnothing(i\neq j)$，且 $E_i(i=1,2,\cdots,s)$ 可测，使得当 $x\in E_i$ 时 $f(x)=a_i$，$a_i(i=1,2,\cdots,s)$ 是常数，则说 $f(x)$ 是 E 上的简单函数.

例 2 可测集上的常值函数是简单函数.

例 3 Dirichlet 函数

$$D(x)=\begin{cases}1, & \text{当 } x \text{ 是}[0,1]\text{中的有理数时;}\\ 0, & \text{当 } x \text{ 是}[0,1]\text{中的无理数时}\end{cases}$$

是[0, 1]上的简单函数.

由定理 5 可知**简单函数都是可测函数**,而由定理 4 可知一列简单函数的极限函数也是可测函数. 反过来还有:

定理 6 **设 $f(x)$ 在 $E\subseteq \mathbf{E}^n$ 上可测,则 $f(x)$ 总可以表示成一列简单函数 $\{\varphi_n(x)\}$ 的极限函数,即 $f(x)=\lim\limits_{n\to\infty}\varphi_n(x)$,而且有 $|\varphi_1(x)|\leqslant|\varphi_2(x)|\leqslant\cdots$.**

证明在这里略去,读者可参阅本篇末的参考文献.

由此得到:

定理 7 **函数 $f(x)$ 在 $E\subseteq \mathbf{E}^n$ 上可测的充分必要条件是 $f(x)$ 总可以表示为一列简单函数 $\{\varphi_n(x)\}$ 的极限函数,其中 $|\varphi_1(x)|\leqslant|\varphi_2(x)|\leqslant\cdots$.**

第二节 可测函数的构造

本节讨论可测函数和连续函数的关系,从而揭示可测函数的结构.

定义 1 **设 $E\subseteq \mathbf{E}^n$, $f(x)$ 是 E 上的实值函数, $x_0\in E$,如果 $\forall\varepsilon>0$, $\exists\delta>0$,使得当 $x\in O(x_0,\delta)\cap E$ 时,有 $f(x)\in O(f(x_0),\varepsilon)$,则说 $f(x)$ 在 x_0 处连续,此时也说 x_0 是 $f(x)$ 的连续点. 若 $f(x)$ 在 E 上每一点都连续,则说 $f(x)$ 在 E 上连续,此时我们说 $f(x)$ 是 E 上的连续函数.**

定理 1 **可测集 $E\subseteq \mathbf{E}^n$ 上的连续函数 $f(x)$ 是可测函数.**

证 $\forall a\in\mathbf{R}$,设 $x\in E[f>a]$,由于 $f(x)$ 是 E 上的连续函数,所以存在 x 的某邻域 $O(x,\delta)$,使得

$$O(x,\delta)\cap E\subseteq E[f>a],$$

因此令 $G=\bigcup\limits_{x\in E[f>a]}O(x,\delta)$,则

$$G\cap E=\bigcup_{x\in E[f>a]}(O(x,\delta)\cap E)\subseteq E[f>a].$$

而另一方面显然有 $E[f>a]\subseteq G$,从而 $E[f>a]\subseteq G\cap E$,于是

$$E[f>a]=G\cap E.$$

而 G 是一簇开集的并集,仍为开集,故可测,而已知 E 可测,所以 $E[f>a]$ 可测,于是结论成立.

定理 2 **(Лузин(鲁津)定理)设 $f(x)$ 是 $E\subseteq \mathbf{E}^n$ 上 a. e. 有限的可测函数,则**

$\forall \delta > 0$，**存在闭集** $F_\delta \subseteq E$，**使得** $m[(E-F_\delta)] < \delta$，**且** $f(x)$**在** F_δ **上连续.**

简言之：在 E **上 a. e. 有限的可测函数是“基本上连续”的函数.**

证 由于 $E_0 = E[|f| = +\infty]$ 是零测度集，而必要时总可以用 $E - E_0$ 代替 E，所以不妨假设 $f(x)$ 在 E 上处处有限.

将 $\mathbf{E}^1$ 中的全体有理数排成一列为

$$r_1, r_2, \cdots, r_n, \cdots,$$

令

$$E_{nk} = E\left[r_n < f(x) \leqslant r_n + \frac{1}{k}\right] \text{（其中}n, k = 1, 2, \cdots\text{）},$$

则 E_{nk} 均可测.

于是对于每个 E_{nk} 都存在开集 $G_{nk} \supseteq E_{nk}$，且

$$m(G_{nk} - E_{nk}) < \frac{\delta}{2^{n+k+1}}.$$

令 $E_\delta = \bigcup\limits_{n, k=1}^{\infty} (G_{nk} - E_{nk})$，则 E_δ 可测且

$$mE_\delta \leqslant \sum_{n, k=1}^{\infty} m(G_{nk} - E_{nk}) < \sum_{n, k=1}^{\infty} \frac{\delta}{2^{n+k+1}} = \frac{\delta}{2}.$$

下面证明 $f(x)$ 在 $E - E_\delta$ 上连续：

$\forall \varepsilon > 0$，取正整数 k_0，使得 $\frac{1}{k_0} < \varepsilon$. $\forall\ x_0 \in E - E_\delta$，总有正整数 n_0，使得

$$f(x_0) \in \left(r_{n_0}, r_{n_0} + \frac{1}{k_0}\right],$$

从而 $x_0 \in E_{n_0k_0} \subseteq G_{n_0k_0}$.

由于 $G_{n_0k_0}$ 是开集，所以存在 $O(x_0, \eta) \subseteq G_{n_0k_0}$，当 $x \in O(x_0, \eta) \cap E_{n_0k_0}$ 时，有

$$|f(x) - f(x_0)| \leqslant \frac{1}{k_0} < \varepsilon.$$

而又由于

$$\begin{aligned} O(x_0, \eta) \cap (E - E_\delta) &= O(x_0, \eta) \cap \left[E - \bigcup_{n, k=1}^{\infty} (G_{nk} - E_{nk})\right] \\ &\subseteq O(x_0, \eta) \cap [E - (G_{n_0k_0} - E_{n_0k_0})] \\ &= O(x_0, \eta) \cap [(E \cap G^c_{n_0k_0}) \cup E_{n_0k_0}] \\ &= O(x_0, \eta) \cap E_{n_0k_0}, \end{aligned}$$

因此当 $x \in O(x_0, \eta) \cap (E - E_\delta)$ 时，就有

$$|f(x) - f(x_0)| < \varepsilon$$

成立，这说明了 $f(x)$ 作为 $E - E_\delta$ 上的函数在 x_0 处连续.

由 $x_0 \in E - E_\delta$ 的任意性可知 $f(x)$ 在 $E - E_\delta$ 上连续.

由于 $E - E_\delta$ 可测，所以存在闭集 $F_\delta \subseteq (E - E_\delta)$，使得

$$m((E - E_\delta) - F_\delta) < \frac{\delta}{2},$$

而 $f(x)$ 在 F_δ 上连续且

$$\begin{aligned} m(E - F_\delta) &= m(E - (E - E_\delta)) + m((E - E_\delta) - F_\delta) \\ &< m(E \cap E_\delta) + \frac{\delta}{2} \leqslant mE_\delta + \frac{\delta}{2} \\ &< \frac{\delta}{2} + \frac{\delta}{2} = \delta. \end{aligned}$$

Лузин 定理揭示了可测函数与连续函数的关系，使我们对于可测函数的结构有了进一步的认识，而且 Лузин 定理的逆定理也成立.

第三节　可测函数列的收敛性

本节讨论可测函数列几乎处处收敛、几乎一致收敛与依测度收敛三者之间的关系.

定义 1　设 $f_n(x)$ $(n = 1, 2, \cdots)$ 与 $f(x)$ 都是可测集 $E \subseteq \mathbf{E}^n$ 上的 a. e. 有限的广义实值函数，

若 $m(E[f_n \nrightarrow f]) = 0$，则说 $\{f_n(x)\}$ 在 E 上几乎处处收敛于 $f(x)$，记为 $f_n(x) \xrightarrow{\text{a.e.}} f(x)$ 于 E.

若 $\forall \delta > 0$，存在 E 的可测子集 E_1，且 $mE_1 < \delta$，使得 $\{f_n(x)\}$ 在 $E_\delta = E - E_1$ 上一致收敛于 $f(x)$，则说 $\{f_n(x)\}$ 在 E 上几乎一致收敛于 $f(x)$，记为 $f_n(x) \xrightarrow{\text{a.u.}} f(x)$ 于 E.

定理 1　设 $f_n(x)$ $(n = 1, 2, \cdots)$ 与 $f(x)$ 都是可测集 $E \subseteq \mathbf{E}^n$ 上的 a. e. 有限的可测函数，则

(1) $f_n(x) \xrightarrow{\text{a.e.}} f(x)$ 于 E 的充分必要条件是 $\forall \varepsilon > 0$，恒有

$$m\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E[|f_n(x) - f(x)| \geqslant \varepsilon]\right) = 0,$$

特别地,当 $mE<+\infty$ 时,$f_n(x)\xrightarrow{\text{a.e.}}f(x)$ 于 E 的充分必要条件是 $\forall\varepsilon>0$,有

$$\lim_{N\to\infty}m\left(\bigcup_{n=N}^{\infty}E[\,|f_n(x)-f(x)|\geqslant\varepsilon]\right)=0;$$

(2) $f_n(x)\xrightarrow{\text{a.u.}}f(x)$ 于 E 的充分必要条件是 $\forall\varepsilon>0$,有

$$\lim_{N\to\infty}m\left(\bigcup_{n=N}^{\infty}E[\,|f_n(x)-f(x)|\geqslant\varepsilon]\right)=0.$$

定理 1 的证明略去,读者可参阅本篇末的参考文献.

定义 2 设 $f_n(x)$ $(n=1,2,\cdots)$ 与 $f(x)$ 都是可测集 $E\subseteq\mathbf{E}^n$ 上的 a.e. 有限的可测函数,若 $\forall\varepsilon>0$,有

$$\lim_{n\to\infty}m(E[\,|f_n(x)-f(x)|\geqslant\varepsilon])=0,$$

则说 $\{f_n(x)\}$ 在 E 上依测度收敛于 $f(x)$,记为 $f_n(x)\Rightarrow f(x)$ 于 E.

定理 2 设 $f_n(x)(n=1,2,\cdots)$ 与 $f(x)$ 都是可测集 $E\subseteq\mathbf{E}^n$ 上的 a.e. 有限的可测函数,

当 $mE\leqslant+\infty$ 时,有

(1) 若 $f_n(x)\xrightarrow{\text{a.u.}}f(x)$ 于 E,则 $f_n(x)\xrightarrow{\text{a.e.}}f(x)$ 于 E;

(2) 若 $f_n(x)\xrightarrow{\text{a.u.}}f(x)$ 于 E,则 $f_n(x)\Rightarrow f(x)$ 于 E;

(3) 若 $f_n(x)\Rightarrow f(x)$ 于 E,则必有 $\{f_n(x)\}$ 的子列 $\{f_{n_k}(x)$,使得 $f_{n_k}(x)\xrightarrow{\text{a.u.}}f(x)$ 于 E;

(4) (F. Riesz 定理)若 $f_n(x)\Rightarrow f(x)$ 于 E,则必有 $\{f_n(x)\}$ 的子列 $\{f_{n_k}(x)\}$,使得 $f_{n_k}(x)\xrightarrow{\text{a.e.}}f(x)$ 于 E;

当 $mE<+\infty$ 时,有

(5) (Егоров(叶果洛夫)定理)若 $f_n(x)\xrightarrow{\text{a.e.}}f(x)$ 于 E,则 $f_n(x)\xrightarrow{\text{a.u.}}f(x)$ 于 E;

(6) (Lebesgue 定理)若 $f_n(x)\xrightarrow{\text{a.e.}}f(x)$ 于 E,则 $f_n(x)\Rightarrow f(x)$ 于 E.

证 由定理 1,可知若 $f_n\xrightarrow{\text{a.u.}}f$ 于 E,则 $\forall\varepsilon>0$,有

$$\lim_{N\to\infty}m\left(\bigcup_{n=N}^{\infty}E[\,|f_n-f|\geqslant\varepsilon]\right)=0.$$

令 $B_N=\bigcup_{n=N}^{\infty}E[\,|f_n-f|\geqslant\varepsilon]$,则 $B_N\supseteq B_{N+1}$ $(N=1,2,\cdots)$,且存在 N_0,使得 $m(B_{N_0})<+\infty$,由测度的上连续性,有

$$\lim_{N\to\infty} m(B_N) = m(\bigcap_{N=1}^{\infty} B_N) = m(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E[\,|f_n - f| \geqslant \varepsilon]) = 0.$$

再由定理 1,可知有 $f_n \xrightarrow{\text{a. e.}} f$ 于 E,即(1) 成立.

因为

$$m(E[\,|f_n - f| \geqslant \varepsilon]) \leqslant m(\bigcup_{n=N}^{\infty} E[\,|f_n - f| \geqslant \varepsilon]) \quad (n \geqslant N),$$

从而

$$\lim_{n\to\infty} m(E[\,|f_n - f| \geqslant \varepsilon]) = 0,$$

于是 $f_n \Rightarrow f$ 于 E, 即(2)成立.

下面证明(3)成立:

由于 $f_n \Rightarrow f$ 于E,则 $\forall \varepsilon > 0$ 与 $\forall k \in \mathbf{Z}^+$,存在正整数 $K = K(k, \varepsilon)$. 使得当 $n \geqslant K$ 时,有

$$m(E[\,|f_n - f| \geqslant \varepsilon]) < \frac{1}{2^{k+1}}.$$

令

$$n_1 = K(1, \varepsilon),\ n_2 = \max\{n_1 + 1,\ K(2, \varepsilon)\},\ \cdots,$$

$$n_k = \max\{n_{k-1} + 1,\ K(k, \varepsilon)\},\ \cdots$$

则
$$n_1 < n_2 < \cdots < n_k < \cdots,$$

且
$$m(E[\,|f_{n_k} - f| \geqslant \varepsilon]) < \frac{1}{2^{k+1}} \quad (k = 1, 2, \cdots).$$

于是, $\forall N \in \mathbf{Z}^+$, 有

$$\begin{aligned} & m(\bigcup_{k=N}^{\infty} E[\,|f_{n_k} - f| \geqslant \varepsilon]) \\ \leqslant & \sum_{k=N}^{\infty} m(E[\,|f_{n_k} - f| \geqslant \varepsilon]) \\ < & \sum_{k=N}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{2^N}, \end{aligned}$$

故 $\lim_{N\to\infty} m(\bigcup_{k=N}^{\infty} E[\,|f_{n_k} - f| \geqslant \varepsilon]) = 0$,即 $f_{n_k} \xrightarrow{\text{a. u.}} f$ 于 E.

由(3)与(1)可立得(4).

由定理 1 可立得(5).

由(5)和(2)可立得(6).

注意 若去掉条件 $mE < +\infty$,则定理的(5),(6)不再成立.

例 1 取 $E=(-\infty,+\infty)$，作函数列 $f_n(x)=\dfrac{x}{n}$，$n\in\mathbf{Z}^+$，则有

$$f_n(x)\xrightarrow{\text{a.e.}}0\ (x\in E),$$

任取 $E_1\subseteq E$，且 $mE_1<+\infty$，可知在 $E-E_1$ 上 $f_n(x)$ 不可能一致收敛于 0，即在 E 上 $f_n(x)\xrightarrow{\text{a.u.}}0$ 不成立.

例 2 取 $E=(0,+\infty)$，作函数列

$$f_n(x)=\begin{cases}1, & \text{当 } x\in(0,n] \text{ 时};\\ 0, & \text{当 } x\in(n,+\infty) \text{ 时},\end{cases}\quad n\in\mathbf{Z}^+.$$

显然 $f_n(x)\xrightarrow{\text{a.e.}}1(x\in E)$，但是当 $0<\varepsilon<1$ 时，

$$E[\,|f_n(x)-1|\geqslant\varepsilon]=(n,+\infty) \text{ 且 } m(n,+\infty)=+\infty,$$

说明 $f_n(x)\not\Rightarrow 1$ 于 E.

定理 2 揭示了几乎处处收敛、几乎一致收敛与依测度收敛三者之间的关系，下面对依测度收敛再介绍一些基本性质.

定理 3 设 $mE<+\infty$，$f_n(x)\Rightarrow f(x)$，$g_n(x)\Rightarrow g(x)(x\in E)$，则有

(1) 若 $f_n(x)\Rightarrow h(x)(x\in E)$，则 $f(x)=h(x)$a. e. 于 E；

(2) 若 $a,b\in\mathbf{R}$，则 $af_n(x)+bg_n(x)\Rightarrow af(x)+bg(x)$ 于 E；

(3) $f_n(x)g_n(x)\Rightarrow f(x)g(x)$ 于 E；

(4) 若 $g_n(x)$和 $g(x)$在 E 上几乎处处不为 0，则 $\dfrac{f_n(x)}{g_n(x)}\Rightarrow\dfrac{f(x)}{g(x)}$ 于 E；

(5) $|f_n(x)|\Rightarrow|f(x)|$ 于 E.

习　题　1-4

1. 讨论集合的特征函数的可测性.

2. 证明：在一个零测集上任意改变一个函数的定义，不影响函数的可测性.

3. 证明 Лузин 定理的逆定理.

4. 若 E 是有界可测集，$f(x)$是 E 上 a. e. 有限的函数，则 $f(x)$可测的充分必要条件是存在一列在整个空间上连续的函数 $\varphi_n(x)$，使 $\varphi_n(x)\xrightarrow{\text{a.e.}}f(x)$于 E，试就一维空间的情形证明.

5. 设 $f(x)$是 E 上几乎处处有限的可测函数，$mE<+\infty$，则 $\forall\varepsilon>0$，存在有界可测函数 $g(x)$，使得 $mE[f(x)\neq g(x)]<\varepsilon$.

6. 设 $f_n(x)\Rightarrow f(x)$ 于 E，且 $f(x)=g(x)$a. e. 于 E，则 $f_n(x)\Rightarrow g(x)$ 于 E.

第五章　Lebesgue 积分

本章将在 Lebesgue 测度基础上建立 Lebesgue 积分，这是实变函数的重点内容. Lebesgue 积分是 20 世纪初法国数学家 Lebesgue 提出来的，它的发展比数学分析中的 Riemann 积分要迟半个世纪. 我们知道 Riemann 积分在求面积、体积、物体重心、矩量等问题中起了重大作用，但 Riemann 积分主要针对于连续函数或者基本上连续的函数，这对于量子力学中的物理量或一般随机变量的数学期望值来说显然是不够用的，至于积分与极限交换顺序，重积分交换顺序，Newton-Leibniz 公式与一些重要结果的运用往往要加上一些很强的条件，而 Lebesgue 积分处理这一类问题是相当自由和灵活的，它在很大程度上摆脱了 Riemann 积分的困境，而且大大地扩充了可积函数的范围，成为现代分析中不可缺少的工具.

第一节　Lebesgue 积分的定义

Lebesgue 积分有很多种定义方法，为了便于和数学分析中学过的 Riemann 积分相对照，我们考虑采用"大小和式"也即"确界式"定义，这样能更好地突出应用积分概念时的基本思想——分割求和.

本节先建立测度有限的集合上有界函数的积分，对于更一般的可积函数，我们放到后面再进行讨论.

定义 1　设 $E_1, E_2, \cdots, E_s$ 是 $\mathbf{E}^n$ 中一列非空可测集，如果

$$E_i \cap E_j = \varnothing (i \neq j) \text{ 且 } E = \bigcup_{i=1}^{s} E_i,$$

则说有限集合族 $E_1, E_2, \cdots, E_s$ 是 E 的一个可测分划，简称分划，记为 $D = \{E_i\}$.

设 $D' = \{E'_i\}$ 是 E 的另一分划，如果对于任一 $E'_j \in D'$，存在 $E_i \in D$，使得 $E'_j \subseteq E_i$，则说 D' 比 D 细.

显然有：

引理 1　给定 E 的任两个分划 D, D'，必然存在比它们都细的第三个分划 $D'' = \{E_i \cap E'_j \mid E_i \in D, E'_j \in D', \text{且 } E_i \cap E'_j \neq \varnothing\}$，记为 $D'' = DD'$.

定义 2　设 $f(x)$ 是定义在测度有限的可测集 $E \subseteq \mathbf{E}^n$ 上的有界函数，对 E 的任一分划 $D = \{E_i\}$，令

$$B_i = \sup_{x \in E_i} f(x),\ b_i = \inf_{x \in E_i} f(x),$$

则 $\sum_i B_i mE_i$，$\sum_i b_i mE_i$ 分别叫做 $f(x)$ 关于分划 D 的大和及小和，并分别记为 $S(D, f)$ 及 $s(D, f)$.

引理 2 (1) 设 $B = \sup_{x \in E} f(x)$，$b = \inf_{x \in E} f(x)$，则

$$bmE \leqslant s(D, f) \leqslant S(D, f) \leqslant BmE;$$

(2) 设分划 D' 比 D 细，则 $s(D, f) \leqslant s(D', f)$，$S(D', f) \leqslant S(D, f)$；

(3) 对于任两个分划 D, D'，总有 $s(D, f) \leqslant S(D', f)$；

(4) $\sup_D s(D, f) \leqslant \inf_D S(D, f)$，这里的上下确界是对 E 的所有可能分划取的.

证 由定义可知(1)成立，而(4)可以由(3)直接推出，故只证明(2)与(3).

关于(2)，设 $D = \{E_i\}$，比它细的 $D' = \{E'_j\}$，则

$$s(D, f) = \sum_i b_i mE_i = \sum_i b_i \Big(\sum_{E'_j \subseteq E_i} mE'_j\Big)$$

$$\leqslant \sum_i \sum_{E'_j \subseteq E_i} b'_j mE'_j = \sum_j b'_j mE'_j = s(D', f),$$

同样可证 $S(D', f) \leqslant S(D, f)$.

关于(3)，对于分划 D, D'，作 $D'' = DD'$，则 D'' 比 D, D' 都细，从而由(1)，(2)有

$$s(D, f) \leqslant s(D'', f) \leqslant S(D'', f) \leqslant S(D', f).$$

定义 3 设 $f(x)$ 是定义在测度有限的可测集 $E \subseteq \mathbf{E}^n$ 上的有界函数，$\overline{\int_E} f(x)\mathrm{d}x = \inf_D S(D, f)$ 与 $\underline{\int_E} f(x)\mathrm{d}x = \sup_D s(D, f)$ 分别叫做 $f(x)$ 在 E 上的 Lebesgue 上、下积分，简称 L 上、下积分.

如果 $\overline{\int_E} f(x)\mathrm{d}x = \underline{\int_E} f(x)\mathrm{d}x$，则说 $f(x)$ 在 E 上 Lebesgue 可积. 简称 L 可积，此时记 $\int_E f(x)\mathrm{d}(x) = \overline{\int_E} f(x)\mathrm{d}x = \underline{\int_E} f(x)\mathrm{d}x$，叫做 $f(x)$ 在 E 上的 Lebesgue 积分，简称 L 积分.

注意 $\mathbf{E}^n$ 中测度有限的可测集上有界函数的 L 积分的定义在形式上同 Riemann 积分(以后简称 R 积分)完全类似，只是采用的测度和分划不同.

定理 1 设 $f(x)$ 是定义在测度有限的可测集 $E \subseteq \mathbf{E}^n$ 上的有界函数，则 $f(x)$ 在 E 上 L 可积的充分必要条件是 $\forall \varepsilon > 0$，存在 E 的分划 D，使得

$$S(D, f) - s(D, f) < \varepsilon.$$

证 充分性：由

$$s(D, f) \leqslant \underline{\int}_E f(x)\mathrm{d}x \leqslant \overline{\int}_E f(x)\mathrm{d}x \leqslant S(D, f)$$

得

$$0 \leqslant \overline{\int}_E f(x)\mathrm{d}x - \underline{\int}_E f(x)\mathrm{d}x \leqslant S(D, f) - s(D, f) < \varepsilon,$$

由 ε 的任意性可知

$$\overline{\int}_E f(x)\mathrm{d}x - \underline{\int}_E f(x)\mathrm{d}x = 0,$$

所以 $f(x)$ 可积.

必要性：设 $\overline{\int}_E f(x)\mathrm{d}x = \underline{\int}_E f(x)\mathrm{d}x$，由上、下积分的定义，$\forall \varepsilon > 0$，存在分划 D_1，D_2，使得

$$S(D_1, f) - \overline{\int}_E f(x)\mathrm{d}x < \frac{\varepsilon}{2}, \quad \underline{\int}_E f(x)\mathrm{d}x - s(D_2, f) < \frac{\varepsilon}{2}.$$

因此，对于分划 $D = D_1D_2$，由引理 2 的(2)有

$$S(D, f) - \overline{\int}_E f(x)\mathrm{d}x < \frac{\varepsilon}{2},$$

$$\underline{\int}_E f(x)\mathrm{d}x - s(D, f) < \frac{\varepsilon}{2},$$

两式相加，即得 $S(D, f) - s(D, f) < \varepsilon$.

在第四章我们引进了可测函数，从下面的定理可以看出可测函数的定义事实上是针对着 L 积分的需要而给出的.

定理 2 **设 $f(x)$ 是定义在测度有限的可测集 $E \subseteq \mathbf{E}^n$ 上的有界函数，则 $f(x)$ 在 E 上 L 可积的充分必要条件是 $f(x)$ 在 E 上可测.**

证 充分性：设 $f(x)$ 在 E 上可测，且 $b < f(x) \leqslant B$，$\forall \delta > 0$，作$[b, B]$的任意分划：

$$b = y_0 < y_1 < \cdots < y_k = B,$$

使得 $\max\limits_i(y_i - y_{i-1}) < \delta$.

令 $E_i = E[y_{i-1} < f(x) \leqslant y_i]$，$i = 1, 2, \cdots, k$，则各 E_i 可测且互不相交，又 $E = \bigcup\limits_{i=1}^{k} E_i$，所以$D = \{E_i\}$ 构成 E 的一个分划，且显然有

$$S(D, f) \leqslant \sum_i y_i mE_i, \ s(D, f) \geqslant \sum_i y_{i-1} mE_i \text{(这是因为 } y_{i-1} \leqslant b_i \leqslant B_i \leqslant y_i\text{)}.$$

从而有

$$S(D, f) - s(D, f) \leqslant \sum (y_i - y_{i-1}) mE_i \leqslant \delta mE,$$

由 δ 的任意性和定理 1 可知 $f(x)$在 E 上 L 可积.

必要性:设 $f(x)$在 E 上 L 可积,我们设法用两列简单函数从上下两方面来逼近 $f(x)$,从而证明 $f(x)$可测.

首先由定理 1,对正数 $\frac{1}{k}$, $k = 1, 2, \cdots$, 依次作分划 D_k,使得

$$S(D_k, f) - s(D_k, f) < \frac{1}{k}.$$

不妨假设这列$\{D_k\}$中的分划一个比一个细(如若不然可用

$$D_1^* = D_1, \ D_2^* = D_1 D_2, \ \cdots$$

代替 $D_1, D_2, \cdots$,此时仍有

$$S(D_k^*, f) - s(D_k^*, f) < \frac{1}{k}).$$

设 $D_k = \{E_1^{(k)}, E_2^{(k)}, \cdots, E_{m_k}^{(k)}\}$,

$$\sup_{x \in E_i^{(k)}} f(x) = B_i^{(k)}, \quad \inf_{x \in E_i^{(k)}} f(x) = b_i^{(k)}, \ i = 1, 2, \cdots, m_k,$$

由此出发作两列简单函数

$$g_k(x) = b_i^{(k)}, \ h_k(x) = B_i^{(k)}, \ x \in E_i^{(k)}, \ i = 1, 2, \cdots, m_k, \ k = 1, 2, \cdots.$$

显然 $g_k(x) \leqslant f(x) \leqslant h_k(x)$, $k = 1, 2, \cdots$, 而且由于 D_k 一个比一个细,就有 $g_k(x)$递增,$h_k(x)$递减,因此

$$\lim_{k\to\infty} g_k(x) = g(x) \text{ 与} \lim_{k\to\infty} h_k(x) = h(x)$$

均存在,且 $g(x) \leqslant f(x) \leqslant h(x)$.

因为 $g(x)$, $h(x)$都是简单函数的极限,故都是可测函数,所以如果可以证明到 $g(x) = h(x)$a. e. 于 E,那么一定就有 $f(x) = g(x) = h(x)$a. e. 于 E,从而就证明了 $f(x)$的可测性.

假设"$g(x) = h(x)$ a. e. 于 E"不成立,则由于

$$E[h(x) - g(x) > 0] = \bigcup_{k=1}^{\infty} E\left[h(x) - g(x) \geqslant \frac{1}{k}\right],$$

必存在 N,使得 $mE\left[h(x)-g(x)\geqslant\frac{1}{N}\right]>0$.

记 $E^*=E\left[h(x)-g(x)\geqslant\frac{1}{N}\right]$,则在 E^* 上既然有 $h(x)-g(x)\geqslant\frac{1}{N}$,那么对于任意的 k 就更有

$$h_k(x)-g_k(x)\geqslant\frac{1}{N},$$

也就是当 $x\in E^*\cap E_i^{(k)}$ 时,$B_i^{(k)}-b_i^{(k)}\geqslant\frac{1}{N}$,这样一来便有

$$\begin{aligned}S(D_k,f)-s(D_k,f)&=\sum_i(B_i^{(k)}-b_i^{(k)})mE_i^{(k)}\\&\geqslant\sum_i(B_i^{(k)}-b_i^{(k)})m(E_i^{(k)}\cap E^*)\\&\geqslant\frac{1}{N}\sum_i m(E_i^{(k)}\cap E^*)=\frac{1}{N}mE^*.\end{aligned}$$

这同 $S(D_k,f)-s(D_k,f)<\frac{1}{k}\to 0(k\to\infty)$ 矛盾,所以 $g(x)=h(x)$ a. e. 于 E.

由定理 2 显然就有:

定理 3 **设** $f(x)$, $g(x)$**是测度有限的可测集** $E\subseteq\mathbf{E}^n$ **上的 L 可积函数,则**

$$f(x)+g(x),\ f(x)g(x),\ \frac{f(x)}{g(x)}(\inf_{x\in E}|g(x)|>0),\ |f(x)|$$

在 E 上都是 L 可积的.

第二节 Lebesgue 积分的性质

本节重点介绍 L 积分的一些初等性质,在本节的定理中 E 是 $\mathbf{E}^n$ 中测度有限的可测集,$f(x)$, $g(x)$是 E 上的有界函数.

定理 (1) **设** $f(x)$, $g(x)$**在** E **上 L 可积,则对于任意常数** α, β,**有** $\alpha f(x)+\beta g(x)$ **在** E **上 L 可积,且**

$$\int_E(\alpha f(x)+\beta g(x))\mathrm{d}x=\alpha\int_E f(x)\mathrm{d}x+\beta\int_E g(x)\mathrm{d}x;$$

(2) **设** $f(x)$**在** E **上 L 可积,则** $f(x)$**在** E **的任意可测子集上 L 可积,且若** $E=E_1\cup E_2$, $E_1\cap E_2=\varnothing$, $E_i(i=1,2)$ **可测,则**

$$\int_E f(x)\mathrm{d}x = \sum_{i=1}^{2}\int_{E_i} f(x)\mathrm{d}x;$$

(3) 设 $f(x)$，$g(x)$在 E 上 L 可积，且 $f(x) \leqslant g(x)$ $(x \in E)$，则

$$\int_E f(x)\mathrm{d}x \leqslant \int_E g(x)\mathrm{d}x,$$

特别地，

(i) 如果 $b \leqslant f(x) \leqslant B$，则 $bmE \leqslant \int_E f(x)\mathrm{d}x \leqslant BmE$；当 $mE = 0$ 时，$\int_E f(x)\mathrm{d}x = 0$；当 $f(x) \equiv c$ 时，$\int_E f(x)\mathrm{d}x = cmE$；

(ii) $\left|\int_E f(x)\mathrm{d}x\right| \leqslant \int_E |f(x)|\mathrm{d}x$；

(4) 设 $f(x)$在 E 上 L 可积，$f(x) \geqslant 0$ 且$\int_E f(x)\mathrm{d}x = 0$，则 $f(x) = 0$ a. e. 于 E；

(5)（L 积分的绝对连续性）设 $f(x)$在 E 上 L 可积，则对于任意可测集 $A \subseteq E$，有

$$\lim_{mA \to 0}\int_A f(x)\mathrm{d}x = 0.$$

证　只证(2)，(4)，(5)，其余读者自证.

(2) 前半部分显然，下面证明后半部分：

由于 $f(x)$在 E_1 和 E_2 上都可积，则 $\forall \varepsilon > 0$，存在 E_1 的分划 D_{E_1} 与 E_2 的分划 D_{E_2}，使得

$$S(D_{E_1}, f) < \int_{E_1} f(x)\mathrm{d}x + \frac{\varepsilon}{2},$$

$$S(D_{E_2}, f) < \int_{E_2} f(x)\mathrm{d}x + \frac{\varepsilon}{2},$$

将两分划合并可得 E 的可测分划 $D_E = D_{E_1} + D_{E_2}$.

因此有

$$\int_E f(x)\mathrm{d}x \leqslant S(D_E, f) = S(D_{E_1}, f) + S(D_{E_2}, f)$$

$$< \int_{E_1} f(x)\mathrm{d}x + \int_{E_2} f(x)\mathrm{d}x + \varepsilon.$$

由 ε 的任意性，有

$$\int_E f(x)\mathrm{d}x \leqslant \int_{E_1} f(x)\mathrm{d}x + \int_{E_2} f(x)\mathrm{d}x,$$

同理可得相反的不等式.

(4) 由于 $E = \bigcup_{n=1}^{\infty} E[f \geqslant \frac{1}{n}] \cup E[f = 0]$，可令

$$E_n = E[f \geqslant \frac{1}{n}],$$

则 E_n 可测.

因为

$$0 = \int_E f(x)\mathrm{d}x = \int_{E_n} f(x)\mathrm{d}x + \int_{E-E_n} f(x)\mathrm{d}x \geqslant \int_{E_n} f(x)\mathrm{d}x \geqslant \frac{mE_n}{n},$$

故 $mE_n = 0$，从而 $mE[f > 0] = 0$，于是 $f(x) = 0$ a. e. 于 E.

(5) 设 $|f(x)| \leqslant M$, $x \in E$，则

$$\left| \int_A f(x)\mathrm{d}x \right| \leqslant \int_A |f(x)| \mathrm{d}x \leqslant M \cdot mA \to 0 (mA \to 0).$$

第三节　Riemann 积分与 Lebesgue 积分的关系

在数学分析中，我们学过有界函数的 Riemann 积分，现在对有界函数又定义了一种新的积分——Lebesgue 积分，自然会想到弄清楚两种积分之间的关系.

先看一个例子.

例　考察[0, 1]上的 Dirichlet 函数

$$D(x) = \begin{cases} 1, & \text{当 } x \text{ 是}[0,1]\text{ 中的有理数时；} \\ 0, & \text{当 } x \text{ 是}[0,1]\text{ 中的无理数时} \end{cases}$$

在[0, 1]上的 R 积分与 L 积分.

利用数学分析中的结论：

$f(x)$在$[a, b]$上 R 可积的充分必要条件是 $\forall \varepsilon > 0$，都可以找到一个分划 $T = \{x_i\}$ $(i = 0, 1, 2, \cdots, n)$，当 $\|T\| \to 0$ 时，有

$$\sum_{i=1}^{n} \omega_i (x_i - x_{i-1}) < \varepsilon,$$

其中 $\omega_i = M_i - m_i$, $M_i = \sup\limits_{x_{i-1} \leqslant x \leqslant x_i} f(x)$, $m_i = \inf\limits_{x_{i-1} \leqslant x \leqslant x_i} f(x)$ $(i = 1, 2, \cdots, n)$.

容易知道 $D(x)$在$[0, 1]$上的 R 积分不存在.

但 $D(x)$是$[0, 1]$上的简单函数,因此在$[0, 1]$上可测,所以 $D(x)$在$[0, 1]$上 L 可积.

这个例子证明了确实有 L 可积但不 R 可积的函数存在,从某种意义上说明 L 积分较 R 积分广. 下面的定理将从理论上证明 L 积分是比 R 积分更为普遍的一种积分.

定理 **若 $f(x)$在$[a, b]$上 R 可积,则 $f(x)$在$[a, b]$上必定 L 可积,且**

$$\int_a^b f(x)\mathrm{d}x = \int_{[a,b]} f(x)\mathrm{d}x.$$

证 由已知条件可知 $f(x)$在$[a, b]$上有界,并且对于$[a, b]$的任意分划

$$T: a = x_0 < x_1 < \cdots < x_n = b,$$

总对应着一个可测分划

$$D = \{[a, x_1], (x_1, x_2], \cdots, (x_{n-1}, b]\},$$

于是

$$s(T, f) \leqslant \underline{\int}_{[a,b]} f(x)\mathrm{d}x$$

$$\leqslant \overline{\int}_{[a,b]} f(x)\mathrm{d}x \leqslant S(T, f),$$

从而

$$\underline{\int}_a^b f(x)\mathrm{d}x = \sup_T s(T,f) \leqslant \underline{\int}_{[a,b]} f(x)\mathrm{d}x$$

$$\leqslant \overline{\int}_{[a,b]} f(x)\mathrm{d}x \leqslant \inf_T S(T,f) = \overline{\int}_a^b f(x)\mathrm{d}x,$$

于是定理结论成立.

今后我们将 $f(x)$在$[a, b]$上的 Lebesgue 积分就写成 $\int_a^b f(x)\mathrm{d}x$.

注意 上面的定理谈的是一维空间的情形,对于高维空间也有类似的结论,即 L 积分是 R 积分的推广.

第四节 一般可测函数的 Lebesgue 积分

在前面几节中我们讨论了测度有限的集合上有界函数的 L 积分,因为只在这时才能保证大、小和数有限,现在我们要就一般的情况进行讨论,去掉"测度有

限集”与“有界函数”这两个条件，使 L 积分的定义能适用于更加广泛的函数类.

首先来看非负可测函数的情形：

设 $f(x)$是可测集 $E\subseteq \mathbf{E}^n$ 上的非负实函数，可将 E 用一列单调增加的测度有限的可测集 E_k 来逼近，即

$$E_1\subseteq E_2\subseteq\cdots,\ \lim_{k\to\infty}E_k=E,$$

同时可将 $f(x)$用一列定义在 E_k 上而函数值随着 k 的增大逐渐增大的有界函数 $f_k(x)$来逼近，即

$$f_1(x)\leqslant f_2(x)\leqslant\cdots,\ \lim_{k\to\infty}f_k(x)=f(x),$$

并且可保证每个 $f_k(x)$在 E_k 上可测(此时必然有 $f(x)$在 E 上可测，这是因为 $E[f>a]=\bigcup\limits_{k=1}^{\infty}E_k[f_k>a]$).

例如可取：

$E_k=E\cap K_k$，其中 $K_k=\{(x_1,x_2,\cdots,x_n)\mid |x_i|\leqslant k,\ i=1,2,\cdots,n\}$，$f_k(x)=[f(x)]_k=\min\{f(x),k\}$.

于是$\int_{E_k}f_k(x)\mathrm{d}x$ 存在，且有

$$0\leqslant\int_{E_1}f_1(x)\mathrm{d}x\leqslant\int_{E_2}f_2(x)\mathrm{d}x\leqslant\cdots,$$

因此 $\lim\limits_{k\to\infty}\int_{E_k}f_k(x)\mathrm{d}x$ 存在或为 $+\infty$.

不仅如此，由后面的 Levi 定理可以证明该极限值不依赖于逼近列的选择而惟一确定. 此时，如果有 $mE<+\infty$ 且 $f(x)$在 E 上有界，则因为

$$\int_E[f(x)]_k\mathrm{d}x=\int_{E_k}[f(x)]_k\mathrm{d}x+\int_{E-E_k}[f(x)]_k\mathrm{d}x,$$

两边对 $k\to\infty$取极限，注意到 $m(E-E_k)\to 0$，可得

$$\int_E f(x)\mathrm{d}x=\lim_{k\to\infty}\int_{E_k}[f(x)]_k\mathrm{d}x.$$

于是对于非负可测函数来说我们在已知的 L 积分概念上作如下推广.

定义 1　设 $f(x)\geqslant 0$ 在可测集 $E\subseteq\mathbf{E}^n$ 上可测，

$$\int_E f(x)\mathrm{d}x=\lim_{k\to\infty}\int_{E_k}[f(x)]_k\mathrm{d}x$$

叫做 $f(x)$在 E 上的 L 积分(其中 E_k 与$[f(x)]_k$ 如上所述).

其次来看一般函数的情形：

令 $f^+(x)=\max\{f(x),0\}$，$f^-(x)=\max\{-f(x),0\}$，则

$$f(x)=f^+(x)-f^-(x).$$

易知 $f(x)$ 在 E 上可测的充分必要条件是 $f^+(x)$ 与 $f^-(x)$ 在 E 上可测，且若 $f(x)$ 在测度有限的可测集 E 上可积，则必有

$$\int_E f(x)\mathrm{d}x=\int_E f^+(x)\mathrm{d}x-\int_E f^-(x)\mathrm{d}x.$$

因此对于一般的可测函数而言，将 L 积分作如下推广.

定义 2 设 $f(x)$ 在可测集 $E\subseteq \mathbf{E}^n$ 上可测，如果 $\int_E f^+(x)\mathrm{d}x$ 与 $\int_E f^-(x)\mathrm{d}x$ 不同时为 $+\infty$，则

$$\int_E f(x)\mathrm{d}x=\int_E f^+(x)\mathrm{d}x-\int_E f^-(x)\mathrm{d}x$$

叫做 $f(x)$ 在 E 上的 L 积分，同时说 $f(x)$ 在 E 上积分确定，特别地，当该积分值有限时，说 $f(x)$ 在 E 上 L 可积.

注意 (1) 非负可测函数都是积分确定的，若积分有限则在此定义下 L 可积；

(2) 测度有限的可测集上的可积函数是此定义下的 L 可积函数.

定理 1 (1) 设 $mE=0$（但 $E\neq\varnothing$），则对于 E 上的任意实函数 $f(x)$ 都有 $\int_E f(x)\mathrm{d}x=0$；

(2) 设 $f(x)$ 在 E 上 L 可积，则 $mE[|f|=+\infty]=0$，即 $f(x)$ 在 E 上 a.e. 有限；

(3) 设 $f(x)$ 在 E 上积分确定，则 $f(x)$ 在 E 的任一可测子集 A 上也积分确定. 又如果 $E=A\cup B$，A 与 B 都可测且 $A\cap B=\varnothing$，则

$$\int_E f(x)\mathrm{d}x=\int_A f(x)\mathrm{d}x+\int_B f(x)\mathrm{d}x;$$

(4) 设 $f(x)$ 在 E 上积分确定，且 $f(x)=g(x)$ a.e. 于 E，则 $g(x)$ 在 E 上也积分确定，且 $\int_E f(x)\mathrm{d}x=\int_E g(x)\mathrm{d}x$；

(5) 设 $f(x)$，$g(x)$ 在 E 上非负可测，则

$$\int_E (f(x)+g(x))\mathrm{d}x=\int_E f(x)\mathrm{d}x+\int_E g(x)\mathrm{d}x;$$

(6) 设 $f(x)$，$g(x)$ 在 E 上积分确定且 $f(x)\leqslant g(x)$，则

$$\int_E f(x)\mathrm{d}x \leqslant \int_E g(x)\mathrm{d}x.$$

以上诸性质只须利用定义 1 与定义 2 及本章第二节定理即可简单地得到证明.

注意 (1) 在一个零测集上改变函数值,既不影响函数的可积性也不影响它的积分值,即使在其上函数无定义也未尝不可.

(2) 本章第二节中的定理对于一般可积函数也成立. 由于 L 积分具有绝对可积性,所以 L 积分是一种绝对收敛的积分,而 R 积分不必为绝对收敛. 当广义 R 积分存在时,未必有相应的 L 积分,因此 L 积分虽然是 R 积分的推广却非 R 广义积分的推广. 但对非负可测函数的广义积分而言,如果 R 可积,则必定 L 可积. 详细的论证在这里略去,读者可参阅本篇末的参考文献.

例 设 $f(x)$ 在 $E=[a,b]$ 上可积,则 $\forall\varepsilon>0$,必存在 E 上的连续函数 $\varphi(x)$,使得$\int_a^b |f(x)-\varphi(x)|\mathrm{d}x<\varepsilon$ 成立.

证 设 $E_n=E[|f|>n]$(n 是正整数),由于 $f(x)$在 E 上几乎处处有限,所以 $mE_n\to 0(n\to\infty)$.

由积分的绝对连续性, $\forall\varepsilon>0$, $\exists N\in\mathbf{Z}^+$, 使得

$$N\cdot mE_N\leqslant\int_{E_N}|f(x)|\mathrm{d}x<\frac{\varepsilon}{4}.$$

在 $E-E_N$ 上对 $f(x)$ 利用 Лузин 定理,可知存在闭集 $F_N\subseteq(E-E_N)$ 与 $\mathbf{E}^1$ 上的连续函数 $\varphi(x)$ 使得

(1) $m((E-E_N)-F_N)<\frac{\varepsilon}{4N}$;

(2) $x\in F_N$ 时, $f(x)=\varphi(x)$,且$\sup\limits_{x\in\mathbf{E}^1}|\varphi(x)|=\sup\limits_{x\in F_N}|f(x)|\leqslant N$.

所以有

$$\int_a^b|f(x)-\varphi(x)|\mathrm{d}x=\int_{E_N}|f(x)-\varphi(x)|\mathrm{d}x+\int_{E-E_N}|f(x)-\varphi(x)|\mathrm{d}x$$

$$\leqslant\int_{E_N}|f(x)|\mathrm{d}x+\int_{E_N}|\varphi(x)|\mathrm{d}x+\int_{(E-E_N)-F_N}|f(x)-\varphi(x)|\mathrm{d}x+0$$

$$\leqslant\frac{\varepsilon}{4}+N\cdot mE_N+2N\cdot\frac{\varepsilon}{4N}<\frac{\varepsilon}{4}+\frac{\varepsilon}{4}+\frac{\varepsilon}{2}=\varepsilon.$$

第五节　逐项积分定理

本节主要讨论积分和极限交换顺序的问题,我们将看到在处理这个问题上

L 积分所要求的条件比 R 积分弱得多,这正是 L 积分的最大成功之处.

一、非负可测函数列的逐项积分定理

定理 1 (Levi 逐项积分定理)设$\{f_k(x)\}$是可测集 $E\subseteq \mathbf{E}^n$ 上一列非负可测函数,且在 E 上有 $f_k(x)\leqslant f_{k+1}(x)$, $k=1,2,\cdots$, $\lim\limits_{k\to\infty} f_k(x)=f(x)$,则

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x=\int_E f(x)\mathrm{d}x.$$

证明略去,读者可参阅本篇末所列参考文献.

定理 2 (Lebesgue 逐项积分定理)设$\{f_k(x)\}$是可测集 $E\subseteq \mathbf{E}^n$ 上一列非负可测函数,则

$$\int_E\Big(\sum_{k=1}^{\infty} f_k(x)\Big)\mathrm{d}x=\sum_{k=1}^{\infty}\int_E f_k(x)\mathrm{d}x.$$

证 设 $g_n(x)=\sum\limits_{k=1}^{n} f_k(x)$, $n=1,2,\cdots$,则$\{g_n(x)\}$是 E 上非负可测递增函数列,由 Levi 定理可得

$$\int_E \lim_{n\to\infty} g_n(x)\mathrm{d}x=\lim_{n\to\infty}\int_E g_n(x)\mathrm{d}x,$$

将 $\lim\limits_{n\to\infty} g_n(x)=\sum\limits_{k=1}^{\infty} f_k(x)$, $\int_E g_n(x)\mathrm{d}x=\sum\limits_{k=1}^{n}\int_E f_k(x)\mathrm{d}x$ 代入上式即可.

定理 3 (Fatou 引理)设$\{f_k(x)\}$是可测集 $E\subseteq \mathbf{E}^n$ 上一列非负可测函数,则

$$\int_E \varliminf_{k\to\infty} f_k(x)\mathrm{d}x\leqslant \varliminf_{k\to\infty}\int_E f_k(x)\mathrm{d}x.$$

证 设 $g_k(x)=\inf\{f_k(x), f_{k+1}(x),\cdots\}$, $k=1,2,\cdots$,则$\{g_k(x)\}$是E上非负可测递增函数列且$\varliminf\limits_{k\to\infty} f_k(x)=\lim\limits_{k\to\infty} g_k(x)$,由 Levi 定理可得

$$\int_E \varliminf_{k\to\infty} f_k(x)\mathrm{d}x=\int_E \lim_{k\to\infty} g_k(x)\mathrm{d}x=\lim_{k\to\infty}\int_E g_k(x)\mathrm{d}x$$

$$=\varliminf_{k\to\infty}\int_E g_k(x)\mathrm{d}x\leqslant \varliminf_{k\to\infty}\int_E f_k(x)\mathrm{d}x(\text{因为 } g_k(x)\leqslant f_k(x)).$$

注意 Fatou 引理中的不等号是可以实现的. 例如,设

$$f_k(x)=\begin{cases} k, & \text{当}\dfrac{1}{2k}\leqslant x\leqslant \dfrac{1}{k}\text{ 时};\\ 0, & \text{当}\dfrac{1}{k}<x\leqslant 1\text{ 或 }0\leqslant x<\dfrac{1}{2k}\text{ 时}.\end{cases}$$

因为 $\lim_{k\to\infty} f_k(x) = 0$，所以$\int_0^1 \lim_{k\to\infty} f_k(x)\mathrm{d}x = 0$.

又有

$$\int_0^1 f_k(x)\mathrm{d}x = \int_0^{\frac{1}{2k}} 0\mathrm{d}x + \int_{\frac{1}{2k}}^{\frac{1}{k}} k\,\mathrm{d}x + \int_{\frac{1}{k}}^1 0\mathrm{d}x = \frac{1}{2},$$

故得

$$\lim_{k\to\infty} \int_0^1 f_k(x)\mathrm{d}x = \frac{1}{2},$$

所以

$$\int_0^1 \lim_{k\to\infty} f_k(x)\mathrm{d}x < \lim_{k\to\infty} \int_0^1 f_k(x)\mathrm{d}x.$$

二、一般可积函数列的逐项积分定理

定理 4 (Vitali 逐项积分定理) 设$\{f_k(x)\}$是测度有限的可测集 $E\subseteq \mathbf{E}^n$ 上一列可积函数，满足

(1) $f_k(x) \Rightarrow f(x)$ 于 E；

(2) $\forall \varepsilon > 0$，$\exists \delta > 0$，使得当 $A \subseteq E$，$mA < \delta$ 时有

$$\left| \int_A f_k(x)\mathrm{d}x \right| < \varepsilon \ (k = 1, 2, \cdots),$$

此时称 $f_k(x)$的积分具有等度绝对连续性，

则 $f(x)$在 E 上 L 可积，且

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

证明较繁，在这里略去，读者可参阅篇末所列参考文献.

定理 5 (Lebesgue 控制收敛定理) 设$\{f_k(x)\}$是可测集 $E\subseteq \mathbf{E}^n$ 上一列可测函数，满足

(1) $f_k(x) \Rightarrow f(x)$ 于 E；

(2) 存在 E 上的非负可积函数 $F(x)$，使得 $|f_k(x)| \leqslant F(x)$ a. e. 于 E ($k = 1, 2, \cdots$)(称$\{f_k(x)\}$ 为 $F(x)$ 所控制，而 $F(x)$ 叫控制函数)，

则 $f(x)$在 E 上 L 可积且

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

证 由于 $f_k(x) \Rightarrow f(x)$，根据 Riesz 定理，存在子列$\{f_{k_i}(x)\}$ a. e. 收敛于 $f(x)$，由 $|f_{k_i}(x)| \leqslant F(x)$a. e. 于 E，得 $|f(x)| \leqslant F(x)$a. e. 于 E.

因为 $F(x)$可积，所以 $f(x)$在 E 上是可积的，同样可得每个 $f_k(x)$在 E 上是可积的.

因为 $F(x)$在 E 上可积，由 L 积分的定义可知，$\forall \varepsilon > 0$，$\exists N > 0$，使得

$$\int_E F(x)\mathrm{d}x < \int_{E_N} [F(x)]_N \mathrm{d}x + \frac{\varepsilon}{5},$$

由此得到

$$\int_{E-E_N} F(x)\mathrm{d}x = \int_E F(x)\mathrm{d}x - \int_{E_N} F(x)\mathrm{d}x < \frac{\varepsilon}{5}.$$

又由积分的绝对连续性可知，存在正数 δ，使得当可测集 $A \subseteq E$ 且 $mA < \delta$ 时，

$$\int_A F(x)\mathrm{d}x < \frac{\varepsilon}{5},$$

对此 δ，又由 $f_k(x) \Rightarrow f(x)$ 可知，存在正数 K，当 $k > K$ 时，

$$mE_N\left[\,|f_k - f| \geqslant \frac{\varepsilon}{5(mE_N+1)}\right] < \delta.$$

令 $H_k = E_N\left[\,|f_k - f| \geqslant \frac{\varepsilon}{5(mE_N+1)}\right]$，则

$$E_N - H_k = E_N\left[\,|f_k - f| < \frac{\varepsilon}{5(mE_N+1)}\right],$$

由于 $E = H_k \cup (E_N - H_k) \cup (E - E_N)$，则当 $k > K$ 时

$$\begin{aligned}
&\left|\int_E f_k(x)\mathrm{d}x - \int_E f(x)\mathrm{d}x\right| = \left|\int_E [f_k(x) - f(x)]\mathrm{d}x\right| \\
\leqslant &\int_{E_N-H_k} |f_k(x) - f(x)|\,\mathrm{d}x + \int_{E-E_N} |f_k(x) - f(x)|\,\mathrm{d}x \\
&+ \int_{H_k} |f_k(x) - f(x)|\,\mathrm{d}x \\
\leqslant &\frac{\varepsilon}{5(mE_N+1)} mE_N + 2\int_{E-E_N} F(x)\mathrm{d}x + 2\int_{H_k} F(x)\mathrm{d}x \\
< &\frac{\varepsilon}{5} + \frac{2\varepsilon}{5} + \frac{2\varepsilon}{5} = \varepsilon.
\end{aligned}$$

推论 1 将上面定理中的条件“$f_k(x) \Rightarrow f(x)$ 于 E”改为“$f_k(x) \xrightarrow{\text{a.e.}} f(x)$ 于 E”，则结论仍然成立.

推论 2 （Lebesgue 有界收敛定理）设 $mE < +\infty$，$\{f_k(x)\}$ 是 $E \subseteq \mathbf{E}^n$ 上一

列几乎处处一致有界的可测函数，若 $f_k(x)\Rightarrow f(x)$ 于 E 或 $f_k(x)\xrightarrow{\text{a. e.}} f(x)$ 于 E，则 $f(x)$ 在 E 上 L 可积且

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x=\int_E f(x)\mathrm{d}x.$$

例 1 计算 $\displaystyle\sum_{n=1}^{\infty}\int_{-1}^{1}\frac{x^2}{(1+x^2)^n}\mathrm{d}x.$

解 令 $f_n(x)=\dfrac{x^2}{(1+x^2)^n}$，则 $f_n(x)$ $(n=1,2,\cdots)$ 在$[-1,1]$上非负连续，从而非负可测. 当 $x=0$ 时，$\displaystyle\sum_{n=1}^{\infty}f_n(x)=0$；当 $x\neq 0$ 时，

$$\sum_{n=1}^{\infty}f_n(x)=\frac{x^2}{1+x^2}\cdot\frac{1}{1-\dfrac{1}{1+x^2}}=1.$$

设 $f(x)=\displaystyle\sum_{n=1}^{\infty}f_n(x)$，则

$$f(x)=\begin{cases}1, & \text{当 } x\neq 0 \text{ 时;}\\ 0, & \text{当 } x=0 \text{ 时.}\end{cases}$$

由 Lebesgue 逐项积分定理，有

$$\sum_{n=1}^{\infty}\int_{-1}^{1}\frac{x^2}{(1+x^2)^n}\mathrm{d}x=\int_{-1}^{1}f(x)\mathrm{d}x=\int_{-1}^{1}\mathrm{d}x=2.$$

例 2 证明：$\displaystyle\lim_{n\to\infty}\int_0^{+\infty}\frac{\ln(x+n)}{n}\mathrm{e}^{-x}\cos x\mathrm{d}x=0.$

证 设 $f_n(x)=\dfrac{\ln(x+n)}{n}\mathrm{e}^{-x}\cos x$，于是有

$$f_n(x)\to 0(n\to\infty).$$

因为

$$\left(\frac{\ln t}{t}\right)'=\frac{1-\ln t}{t^2}<0(t\geqslant 3),$$

所以 $n\geqslant 3$，$x\geqslant 0$ 时，

$$\frac{\ln(x+n)}{n}=\frac{(n+x)}{n}\frac{\ln(x+n)}{(x+n)}$$

$$\leqslant\frac{n+x}{n}\frac{\ln 3}{3}<\frac{\ln 3}{3}(1+x),$$

从而 $|f_n(x)|<\frac{\ln 3}{3}(1+x)\mathrm{e}^{-x}$.

令 $F(x)=\frac{\ln 3}{3}(1+x)\mathrm{e}^{-x}$，则 $F(x)$ 在$[0,+\infty)$上R可积，因此 $F(x)$在$[0,+\infty)$上 L 可积. 由 Lebesgue 控制收敛定理有

$$\lim_{n\to\infty}\int_0^{+\infty}\frac{\ln(x+n)}{n}\mathrm{e}^{-x}\cos x\mathrm{d}x=\int_0^{+\infty}\left(\lim_{n\to\infty}\frac{\ln(x+n)}{n}\mathrm{e}^{-x}\cos x\right)\mathrm{d}x$$

$$=0.$$

第六节 Fubini 定理

本节主要讨论重积分和累次积分的关系以及累次积分中交换积分顺序的问题. 我们将会看到在重积分化为累次积分以及累次积分交换积分顺序的问题上 L 积分所要求的条件比 R 积分弱得多，这也是 L 积分的一个成功之处.

为了建立 Fubini 定理我们要先讨论乘积空间的测度问题.

定义 1 **设 E 是 $\mathbf{E}^p\times\mathbf{E}^q$ 中一点集，x_0 是 $\mathbf{E}^p$ 中一固定点，则 $\mathbf{E}^q$ 中点集**

$$\{y\in\mathbf{E}^q\mid(x_0,\ y)\in E\}$$

叫做 E 关于 x_0 的截面，记为 E_{x_0}.

定理 1 **（截面定理）设 E 是 $\mathbf{E}^p\times\mathbf{E}^q$ 中的可测集，则**

(1) 对于 $\mathbf{E}^p$ 中几乎所有的 x，E_x 是 $\mathbf{E}^q$ 中的可测集，于是自然有 mE_x 在 $\mathbf{E}^p$ 中几乎处处有定义；

(2) mE_x 是 $\mathbf{E}^p$ 中可测函数，且 $mE=\int_{\mathbf{E}^p}mE_x\mathrm{d}x$.

定理 2 **设 A, B 分别是 $\mathbf{E}^p$, $\mathbf{E}^q$ 中的可测集，则 $A\times B$ 是 $\mathbf{E}^p\times\mathbf{E}^q$ 中的可测集，且**

$$m(A\times B)=mA\cdot mB.$$

定理 1 与定理 2 的证明较繁，在这里略去，读者可参阅本篇末所列的参考文献.

定义 2 **设 $f(x)$是 $E\subseteq\mathbf{E}^n$ 上的非负函数，则 $\mathbf{E}^{n+1}$ 中的点集**

$$\{(x,\ z)\mid x\in E,\ 0\leqslant z<f(x)\}$$

叫做 $f(x)$在 E 上的下方图形，记为 $G(E,\ f)$.

定理 3 **（非负可测函数积分的几何意义）设 $f(x)$是可测集 $E\subseteq\mathbf{E}^n$ 上的非负函数，则**

(1) $f(x)$在 E 上可测的充分必要条件是 $G(E,\ f)$是 $\mathbf{E}^{n+1}$ 中的可测集；

(2) **当** $f(x)$**在** E **上可测时**，$\int_E f(x)\mathrm{d}x = mG(E, f)$.

证 当 $f(x)$是非负常值函数时，设 $f(x) = c$, $c \geqslant 0$, c 是常数，则

$$G(E, f) = \begin{cases} E \times [0, c), & \text{当 } c > 0 \text{ 时;} \\ \varnothing, & \text{当 } c = 0 \text{ 时.} \end{cases}$$

所以 $G(E, f)$是 $\mathbf{E}^{n+1}$ 中的可测集.

当 $f(x)$是简单函数时，对于 $E = \bigcup_k E_k$（E_k 为互不相交的可测集）总有

$$G(E, f) = \bigcup_k G(E_k, f),$$

故 $G(E, f)$可测.

当 $f(x)$是非负可测函数时，则存在一列简单函数

$$0 \leqslant \varphi_1(x) \leqslant \varphi_2(x) \leqslant \cdots,$$

使得 $\lim\limits_{n\to\infty}\varphi_n(x) = f(x)$, $x \in E$.

于是有

$$G(E, \varphi_1) \subseteq G(E, \varphi_2) \subseteq \cdots, \text{且} \bigcup_{n=1}^{\infty} G(E, \varphi_n) = G(E, f),$$

而各 $G(E, \varphi_n)$都可测，故 $G(E, f)$可测.

反之，如果 $G(E, f)$可测，由截面定理可知，$mG_x(E, f)$是 $\mathbf{E}^n$ 中几乎处处有定义的可测函数，且

$$mG_x(E, f) = \begin{cases} f(x), & \text{当 } x \in E \text{ 时;} \\ 0, & \text{当 } x \notin E \text{ 时.} \end{cases}$$

所以 $f(x)$在 E 上可测且

$$\int_E f(x)\mathrm{d}x = mG(E, f).$$

推论 1 **设** $f(x)$**是** $E \subseteq \mathbf{E}^n$ **上的可积函数，则**

$$\int_E f(x)\mathrm{d}x = mG(E, f^+) - mG(E, f^-).$$

推论 2 **可测函数** $f(x)$**在** $E \subseteq \mathbf{E}^n$ **上可积的充分必要条件是**$mG(E, f^+)$**与** $mG(E, f^-)$ **都是有限的.**

由上面的定理可以导出 Fubini 定理，它揭示了高维积分与低维积分之间的联系，也是数学分析中重积分化累次积分的推广.

定理 4 (Fubini 定理) $A \subseteq \mathbf{E}^p$, $B \subseteq \mathbf{E}^q$ **是两个可测集，**

(1) **设** $f(P) = f(x, y)$ **在** $A \times B$ **上非负可测，则对于** a. e. **的** $x \in A$, $f(x,$

y）看作 y 的函数在 B 上可测，且

$$\int_{A\times B} f(P)\mathrm{d}P = \int_A \mathrm{d}x \int_B f(x, y)\mathrm{d}y;$$

（2）设 $f(P) = f(x, y)$ 在 $A\times B$ 上可积，则对于 a. e. 的 $x\in A$，$f(x, y)$ 看作 y 的函数在 B 上可积，$\int_B f(x, y)\mathrm{d}y$ 看作 x 的函数在 A 上可积，且

$$\int_{A\times B} f(P)\mathrm{d}P = \int_A \mathrm{d}x \int_B f(x, y)\mathrm{d}y.$$

证　（1）由定理 3，$G(A\times B, f)$ 是 $\mathbf{E}^{p+q+1}$ 中的可测集，且

$$mG(A\times B, f) = \int_{A\times B} f(P)\mathrm{d}P,$$

而又由截面定理可得

$$mG(A\times B, f) = \int_{\mathbf{E}^p} mG(A\times B, f)_x \mathrm{d}x.$$

因为

$$\mathbf{E}^{q+1} \supseteq G(A\times B, f)_x = \begin{cases} \{(y, z) \mid y\in B, 0\leqslant z < f(x, y)\}, & \text{当 } x\in A \text{ 时}; \\ \varnothing, & \text{当 } x\notin A \text{ 时}, \end{cases}$$

所以对于 $x\in A$，截面 $G(A\times B, f)_x$ 实际上就是将 x 固定后，$f(x, y)$ 看作是 y 的函数时，在 B 上的下方图形 $G(B, f_{(x固定)})$. 于是当该截面可测时，由定理 3 有

$$mG(A\times B, f)_x = mG(B, f_{(x固定)}) = \int_B f(x, y)\mathrm{d}y,$$

于是就得到

$$\int_{A\times B} f(P)\mathrm{d}P = \int_A \mathrm{d}x \int_B f(x, y)\mathrm{d}y.$$

（2）设 $f(P)$ 在 $A\times B$ 上可积，则 $f^+(P)$，$f^-(P)$ 在 $A\times B$ 上也可积，由（1）中的等式可得

$$\begin{aligned}\int_{A\times B} f(P)\mathrm{d}P &= \int_{A\times B} f^+(P)\mathrm{d}P - \int_{A\times B} f^-(P)\mathrm{d}P \\ &= \int_A \mathrm{d}x \int_B f^+(x, y)\mathrm{d}y - \int_A \mathrm{d}x \int_B f^-(x, y)\mathrm{d}y \\ &= \int_A \mathrm{d}x \left(\int_B f^+(x, y)\mathrm{d}y - \int_B f^-(x, y)\mathrm{d}y\right) \\ &= \int_A \mathrm{d}x \int_B (f^+(x, y) - f^-(x, y))\mathrm{d}y \\ &= \int_A \mathrm{d}x \int_B f(x, y)\mathrm{d}y.\end{aligned}$$

从 Fubini 定理我们可以看到，只要重积分有限，它就和两个累次积分相等.

例 设 $f(x, y)=\dfrac{x^2-y^2}{(x^2+y^2)^2}$ 定义在 $E=(0, 1)\times(0, 1)$ 上，则可以算出

$$\int_{(0,1)}\mathrm{d}x\int_{(0,1)}f(x, y)\mathrm{d}y=\frac{\pi}{4} \text{ 与} \int_{(0,1)}\mathrm{d}y\int_{(0,1)}f(x, y)\mathrm{d}x=-\frac{\pi}{4},$$

由 Fubini 定理可知 $f(x, y)$ 在 E 上不可积.

第七节 Lebesgue 不定积分

本节主要揭示 $\mathbf{E}^1$ 空间中 Lebesgue 积分与微分之间的关系. 为此先引进有界变差函数和绝对连续函数等概念.

定义 1 **设 $f(x)$ 是 $[a, b]$ 上的有限函数，如果对于 $[a, b]$ 的任意分划 T，$\left\{\sum\limits_{i=1}^{n}|f(x_i)-f(x_{i-1})|\right\}$ 都是一个有界数集，则说 $f(x)$ 是 $[a, b]$ 上的有界变差函数(或囿变函数)，并说 $\sup\limits_{T}\left\{\sum\limits_{i=1}^{n}|f(x_i)-f(x_{i-1})|\right\}$ 是 $f(x)$ 在 $[a, b]$ 上的全变差，记为 $\overset{b}{\underset{a}{\mathrm{V}}}(f)$，当分划为 T 时，$\sum\limits_{i=1}^{n}|f(x_i)-f(x_{i-1})|$ 叫做 $f(x)$ 在 T 下的变差.**

注意 若 $f(x)$ 在有限闭区间 $[a, b]$ 上满足 Lipschitz(利普希茨)条件(存在正数 L，当 $x_1, x_2\in[a, b]$ 时，有 $|f(x_1)-f(x_2)|\leqslant L|x_1-x_2|$)，则 $f(x)$ 是 $[a, b]$ 上的有界变差函数；另外在有限闭区间上的单调有限函数也是有界变差函数，因此有界变差函数不一定是连续函数，反过来，连续函数也不一定是有界变差函数，例如

$$f(x)=\begin{cases} x\cos\dfrac{\pi}{2x}, & \text{当 } 0<x\leqslant 1 \text{ 时；}\\ 0, & \text{当 } x=0 \text{ 时}\end{cases}$$

是 $[0, 1]$ 上的连续函数，但却不是 $[0, 1]$ 上的有界变差函数.

事实上，对 $[0, 1]$ 取分划 T：$0<\dfrac{1}{2n}<\dfrac{1}{2n-1}<\cdots<\dfrac{1}{3}<\dfrac{1}{2}<1$，

容易得到

$$\sum_{i=1}^{2n}|f(x_i)-f(x_{i-1})|=\sum_{i=1}^{n}\frac{1}{i},$$

于是 $\overset{1}{\underset{0}{\mathrm{V}}}(f)=+\infty$.

定理 1 有界变差函数具有下面一些性质：

(1) $f(x)$是$[a, b]$上的有界变差函数，则 $f(x)$在$[a, b]$上有界；

(2) 设 $f(x)$，$g(x)$均是$[a, b]$上的有界变差函数，则对于任意常数 α 与β，有 $\alpha f(x)+\beta g(x)$ 是$[a, b]$上的有界变差函数，且

$$\bigvee_a^b(\alpha f+\beta g)\leqslant|\alpha|\bigvee_a^b(f)+|\beta|\bigvee_a^b(g);$$

(3) 设 $f(x)$，$g(x)$均是$[a, b]$上的有界变差函数，则 $f(x)g(x)$，$\dfrac{f(x)}{g(x)}$ $(|g(x)|>0)$ 是$[a, b]$上的有界变差函数；

(4) 若 $f(x)$是$[a, b]$上的有界变差函数且 $\bigvee\limits_a^b(f)=0$，则 $f(x)$ 必是常数；

(5) 若 $f(x)$是$[a, b]$上的有界变差函数，则 $\forall c\in(a, b)$，有

$$\bigvee_a^c(f)+\bigvee_c^b(f)=\bigvee_a^b(f);$$

(6) 若$\{g_n(x)\}$是$[a, b]$上一列有界变差函数，$\{\bigvee\limits_a^b(g_n)\}$ 有界，且$\lim\limits_{n\to\infty}g_n(x)=g(x)$，$x\in[a, b]$，则 $g(x)$ 是$[a, b]$ 上的有界变差函数，且

$$\bigvee_a^b(g)\leqslant\sup_n\bigvee_a^b(g_n).$$

证明留给读者.

前面提到闭区间上的有限增函数是有界变差函数，由定理 1 可知，两个有限增函数的差还是有界变差函数. 反之，有下面的重要结论：

定理 2 (Jordan 分解定理)在闭区间$[a, b]$上的有界变差函数 $f(x)$可以表示为两个增函数之差.

证 作函数

$$f_1(x)=\frac{1}{2}\left(\bigvee_a^x(f)+f(x)+|f(a)|\right),$$

$$f_2(x)=\frac{1}{2}\left(\bigvee_a^x(f)-f(x)+|f(a)|\right).$$

任取 $x_1, x_2\in[a, b]$，当 $x_1<x_2$ 时，由于

$$|f(x_1)-f(x_2)|\leqslant\bigvee_{x_1}^{x_2}(f)=\bigvee_a^{x_2}(f)-\bigvee_a^{x_1}(f),$$

所以

$$\bigvee_a^{x_1}(f)+f(x_1)\leqslant\bigvee_a^{x_2}(f)+f(x_2),$$

$$\bigvee_a^{x_1}(f)-f(x_1)\leqslant\bigvee_a^{x_2}(f)-f(x_2).$$

这说明 $f_1(x)$ 和 $f_2(x)$ 都是 $[a, b]$ 上的增函数，由 $f_1(x)$ 和 $f_2(x)$ 的定义可知

$$f(x) = \frac{1}{2}(\overset{x}{\underset{a}{\mathrm{V}}}(f) + f(x) + |f(a)|) - \frac{1}{2}(\overset{x}{\underset{a}{\mathrm{V}}}(f) - f(x) + |f(a)|)$$
$$= f_1(x) - f_2(x).$$

此外还有：

定理 3 （Lebesgue 定理）设 $f(x)$ 是 $[a, b]$ 上的有界变差函数，则

(1) $f(x)$ 在 $[a, b]$ 上几乎处处存在导数 $f'(x)$；

(2) $f'(x)$ 在 $[a, b]$ 上可积；

(3) 如果 $f(x)$ 是增函数，有 $\int_a^b f'(x)\mathrm{d}x \leqslant f(b) - f(a)$.

证明略去，读者可参阅本篇末所列参考文献.

定义 2 设 $F(x)$ 是 $[a, b]$ 上的有限函数，如果 $\forall \varepsilon > 0$，$\exists \delta > 0$，使得对于 $[a, b]$ 中互不相交的任意有限个开区间 (a_i, b_i)，$i = 1, 2, \cdots, n$，只要

$$\sum_{i=1}^{n}(b_i - a_i) < \delta,$$

就有

$$\sum_{i=1}^{n} | F(b_i) - F(a_i) | < \varepsilon,$$

则称 $F(x)$ 是 $[a, b]$ 上的绝对连续函数.

不难证明：

绝对连续函数的线性组合，绝对连续函数的乘积，满足 Lipschitz 条件的函数都是绝对连续函数；绝对连续函数是一致连续函数，并且也是有界变差函数.

定义 3 设 $f(x)$ 在 $[a, b]$ 上 L 可积，则 $[a, b]$ 上的函数

$$F(x) = \int_a^x f(t)\mathrm{d}t + C \text{（}C\text{ 为任意常数）}$$

叫做 $f(x)$ 的一个 Lebesgue 不定积分，简称不定积分.

显然有：

定理 4 设 $f(x)$ 在 $[a, b]$ 上 L 可积，则其不定积分是绝对连续函数.

定理 5 设 $F(x)$ 是 $[a, b]$ 上的绝对连续函数，则几乎处处有定义的 $F'(x)$ 在 $[a, b]$ 上 L 可积，且

$$\int_a^x F'(t)\mathrm{d}t = F(x) - F(a).$$

证 由定理 3 和绝对连续函数的性质可知定理前半部分的正确性，下面证明定理后半部分中的等式成立.

在 $(b, b+1]$ 上补充定义 $F(x)$ 的值为常数 $F(b)$，则 $F(x)$ 便成为$[a, b+1]$上的绝对连续函数. 相应于每个正整数 n，作$[a, b]$上的函数

$$\varphi_n(x) = \frac{F\left(x+\frac{1}{n}\right)-F(x)}{\frac{1}{n}},$$

显然 $\varphi_n(x)$在$[a, b]$上连续，从而 $\varphi_n(x)$在$[a, b]$上不仅 L 可积，而且 R 可积. 又因为 $\varphi_n(x)\xrightarrow{\text{a. e.}}F'(x)$于$[a, b]$，所以如果 $\varphi_n(x)$的积分具有等度绝对连续性，由 Vitali 逐项积分定理可知

$$\begin{aligned}\int_a^x F'(t)\mathrm{d}t &= \lim_{n\to\infty}\int_a^x \varphi_n(t)\mathrm{d}t = \lim_{n\to\infty} n\int_a^x [F(t+\frac{1}{n}) - F(t)]\mathrm{d}t \\ &= \lim_{n\to\infty} n[\int_{a+\frac{1}{n}}^{x+\frac{1}{n}} F(t)\mathrm{d}t - \int_a^x F(t)\mathrm{d}t] \\ &= \lim_{n\to\infty} n[\int_x^{x+\frac{1}{n}} F(t)\mathrm{d}t - \int_a^{a+\frac{1}{n}} F(t)\mathrm{d}t] \\ &= F(x) - F(a).\end{aligned}$$

下面证明 $\varphi_n(x)$的积分具有等度绝对连续性，即要证明 $\forall\varepsilon>0$，$\exists\delta>0$，使得当 $A\subseteq[a, b]$，$mA<\delta$ 时，有

$$\left|\int_A \varphi_n(x)\mathrm{d}x\right| < \varepsilon \ (n=1, 2, \cdots).$$

(1) 当 A 为开集时，设 $A=\bigcup_{i=1}^{\infty} I_i$，其中 $I_i=(a_i, b_i)$ $(i=1, 2, \cdots)$ 为 A 的构成区间. 对于任意给定的正整数 n 及每个正整数 k，有

$$\begin{aligned}\int_{\bigcup_{i=1}^{k} I_i} \varphi_n(x)\mathrm{d}x &= \sum_{i=1}^{k}\int_{a_i}^{b_i} \frac{F\left(x+\frac{1}{n}\right)-F(x)}{\frac{1}{n}}\mathrm{d}x \\ &= n\sum_{i=1}^{k}[\int_{b_i}^{b_i+\frac{1}{n}} F(x)\mathrm{d}x - \int_{a_i}^{a_i+\frac{1}{n}} F(x)\mathrm{d}x] \\ &= n\sum_{i=1}^{k}\int_0^{\frac{1}{n}} [F(b_i+x) - F(a_i+x)]\mathrm{d}x \\ &= n\int_0^{\frac{1}{n}} \sum_{i=1}^{k}[F(b_i+x) - F(a_i+x)]\mathrm{d}x.\end{aligned}$$

由于 $F(x)$ 在 $[a, b+1]$ 上绝对连续，于是 $\forall \varepsilon > 0$，$\exists \delta > 0$，使得对于 $[a, b+1]$ 中的任意有限个互不相交的开区间 $\{(c_i, d_i)\}$，当 $\sum_i (d_i - c_i) < \delta$ 时，有

$$\sum_i |F(d_i) - F(c_i)| < \varepsilon.$$

从而当 $mA = \sum_{i=1}^{\infty}(b_i - a_i) < \delta$ 时，$\forall x \in \left[0, \frac{1}{n}\right]$，有

$$\sum_{i=1}^{\infty}[(b_i + x) - (a_i + x)] = \sum_{i=1}^{\infty}(b_i - a_i) < \delta,$$

即对于每个正整数 k，有 $\sum_{i=1}^{k} |F(b_i + x) - F(a_i + x)| < \varepsilon$.

所以

$$\left| \int_{\bigcup_{i=1}^{k} I_i} \varphi_n(x)\mathrm{d}x \right| \leqslant n\int_0^{\frac{1}{n}} \left| \sum_{i=1}^{k}[F(b_i + x) - F(a_i + x)] \right| \mathrm{d}x < \varepsilon.$$

由于

$$|\chi_{\bigcup_{i=1}^{k} I_i}(x)\varphi_n(x)| \leqslant |\varphi_n(x)| \ (x \in [a, b], k = 1, 2, \cdots),$$

$$\lim_{k\to\infty} \chi_{\bigcup_{i=1}^{k} I_i}(x)\varphi_n(x) = \chi_A(x)\varphi_n(x),$$

根据 Lebesgue 控制收敛定理可得

$$\begin{aligned} \left| \int_A \varphi_n(x)\mathrm{d}x \right| &= \left| \int_a^b \chi_A(x)\varphi_n(x)\mathrm{d}x \right| \\ &= \lim_{k\to\infty} \left| \int_a^b \chi_{\bigcup_{i=1}^{k} I_i}(x)\varphi_n(x)\mathrm{d}x \right| \\ &= \lim_{k\to\infty} \left| \int_{\bigcup_{i=1}^{k} I_i} \varphi_n(x)\mathrm{d}x \right| \\ &< \varepsilon. \end{aligned}$$

故

$$\left| \int_A \varphi_n(x)\mathrm{d}x \right| < \varepsilon \ (n = 1, 2, \cdots).$$

(2) 当 A 为 G_δ 型集，$mA < \delta$ 时，设 $A = \bigcap_{i=1}^{\infty} G_i$，$G_i$ 为开集，且 $[a, b] \supseteq G_1 \supseteq G_2 \supseteq \cdots$，由于 $\lim_{i\to\infty} mG_i = mA < \delta$，所以当 i 充分大后，有 $mG_i < \delta$，同(1)，根据 Lebesgue 控制收敛定理可得

$$\left|\int_A \varphi_n(x)\mathrm{d}x\right| = \left|\int_a^b \chi_A(x)\varphi_n(x)\mathrm{d}x\right|$$
$$= \left|\int_a^b \lim_{i\to\infty}\chi_{G_i}(x)\varphi_n(x)\mathrm{d}x\right|$$
$$= \lim_{i\to\infty}\left|\int_a^b \chi_{G_i}(x)\varphi_n(x)\mathrm{d}x\right|$$
$$= \lim_{i\to\infty}\left|\int_{G_i}\varphi_n(x)\mathrm{d}x\right| < \varepsilon(n=1,\ 2,\ \cdots).$$

(3) 当 A 为任意可测集，$mA<\delta$ 时，必有 G_δ 型集 $B\supseteq A$，使得 $m(B-A)=0$，于是

$$\left|\int_A \varphi_n(x)\mathrm{d}x\right| = \left|\int_A \varphi_n(x)\mathrm{d}x + \int_{B-A}\varphi_n(x)\mathrm{d}x\right|$$
$$= \left|\int_B \varphi_n(x)\mathrm{d}x\right| < \varepsilon(n=1,\ 2,\ \cdots).$$

由此我们得到：$F(x)$是$[a,\ b]$上的绝对连续函数的充分必要条件是 $F(x)$是一个可积函数的不定积分，因此绝对连续性是 Lebesgue 积分的特征性质. 值得注意的是要使公式 $\int_a^x F'(t)\mathrm{d}t = F(x)-F(a)$ 成立，一般说来 $F(x)$的绝对连续性是必需的.

最后介绍 L 积分的分部积分法.

定理 6 **(分部积分法)设** $F(x)$**在**$[a,\ b]$**上绝对连续,**$g(x)$**在**$[a,\ b]$**上 L 可积,**$G(x)$**是** $g(x)$**的不定积分,则**

$$\int_a^b F(x)g(x)\mathrm{d}x = F(x)G(x)\big|_a^b - \int_a^b F'(x)G(x)\mathrm{d}x.$$

证 因为 $F(x)$，$G(x)$都在$[a,\ b]$上绝对连续，从而它们都在$[a,\ b]$上有界，即存在常数 $M>0$，$L>0$ 使得对于一切 $x\in[a,\ b]$，有 $|F(x)|\leqslant M$，$|G(x)|\leqslant L$. 由于 $M|g(x)|$ 与 $L|F'(x)|$ 都在$[a,\ b]$ 上可积，而且有

$$|F(x)g(x)|\leqslant M|g(x)|,\quad |F'(x)G(x)|\leqslant L|F'(x)|,$$

所以 $F(x)g(x)$与 $F'(x)G(x)$都在$[a,\ b]$上可积，从而 $F'(x)G(x)+F(x)g(x)$ 在$[a,\ b]$ 上可积.

令 $H(x)=F(x)G(x)$，则 $H(x)$ 在$[a,\ b]$ 上绝对连续，且

$$H'(x) = F'(x)G(x)+F(x)g(x)$$

在$[a,\ b]$上几乎处处有定义，由定理 5 得

$$\int_a^b F'(x)G(x)\mathrm{d}x+\int_a^b F(x)g(x)\mathrm{d}x = H(b)-H(a) = F(x)G(x)\mid_a^b,$$

故
$$\int_a^b F(x)g(x)\mathrm{d}x = F(x)G(x)\mid_a^b - \int_a^b F'(x)G(x)\mathrm{d}x.$$

第八节　Lebesgue-Stieltjes 测度与积分

本节主要简单介绍 L 测度、L 积分的推广，即所谓 Lebesgue-Stieltjes 测度与建立在其上的 Lebesgue-Stieltjes 积分.

设 $\alpha(x)$是定义在 $\mathbf{E}^1$ 上的有限增函数，对任意开区间 $I=(x,x')$，称 $\alpha(x')-\alpha(x)$ 叫做区间 I 的“权”，记为 $|I|=\alpha(x')-\alpha(x)$.

定义 1　(L-S 外测度)对于任意点集 $E\subseteq\mathbf{E}^1$，非负实数 $\inf\limits_{E\subseteq\bigcup\limits_{i=1}^{\infty}I_i}\sum\limits_{i=1}^{\infty}|I_i|$ 叫做关于函数 $\alpha(x)$ 的L-S 外测度，记为 $m_\alpha^* E$.

显然，当 $\alpha(x)=x$ 时，L-S 外测度便成为 L 外测度.

L-S 外测度与 L 外测度有同样的基本性质：

(1) $m_\alpha^* E\geqslant 0$，且 $m_\alpha^*\varnothing=0$　　(非负性)；

(2) 设 $A\subseteq B$，则 $m_\alpha^* A\leqslant m_\alpha^* B$　　(单调性)；

(3) $m_\alpha^*(\bigcup\limits_{i=1}^{\infty}E_i)\leqslant\sum\limits_{i=1}^{\infty}m_\alpha^* E_i$　　(次可数可加性)；

(4) 若 $d(A,B)>0$，则 $m_\alpha^*(A\cup B)=m_\alpha^* A+m_\alpha^* B$.

但在 L 外测度中，区间不论开闭或半开闭，其测度都是区间长度，而在一般的 L-S 外测度中则不然.

定理　(1) $m_\alpha^*(a,b)=\alpha(b-0)-\alpha(a+0)$；

(2) $m_\alpha^*(a,b]=\alpha(b+0)-\alpha(a+0)$；

(3) $m_\alpha^*[a,b]=\alpha(b+0)-\alpha(a-0)$；

(4) $m_\alpha^*[a,b)=\alpha(b-0)-\alpha(a-0)$.

由定理看出，对 $\alpha(x)$取常值的任意开区间 I，总有 $m_\alpha^* I=0$，而对于 $\alpha(x)$ 的任意不连续点 x_0，则有 $m_\alpha^*\{x_0\}=\alpha(x_0+0)-\alpha(x_0-0)>0$，这恰好与 L 外测度情形相反.

注意　改变 $\alpha(x)$在不连续点处的函数值不影响 L-S 外测度的值，因此，有时可要求 $\alpha(x)$为右连续的增函数，从而有 $m_\alpha^*(a,b]=\alpha(b)-\alpha(a)$.

有了 L-S 外测度，便可以仿照由 L 外测度定义 L 可测集与测度的方法，定义 $\mathbf{E}^1$ 中关于 $\alpha(x)$的 L-S 可测集与测度.

定义 2　(L-S 可测集与测度)设 $E\subseteq\mathbf{E}^1$，如果对于任意 $T\subseteq\mathbf{E}^1$，总有

$$m_\alpha^* T = m_\alpha^* (T \cap E) + m_\alpha^* (T \cap E^c),$$

则说 E 是关于 $\alpha(x)$ 的L-S可测集，而 $m_\alpha^* E$ 叫做 E 关于 $\alpha(x)$ 的L-S测度，记为 $m_\alpha E$.

类似于L测度，不难证明：

(1) 任意区间都是L-S可测的；

(2) L-S可测集及其测度具有完全同于 L 可测集及其测度的一切重要性质，例如关于 $\alpha(x)$ 的L-S可测集的并交补运算都是封闭的，有可数可加性等等；

(3) 当 $m_\alpha^* E = 0$ 时，必有 $m_\alpha E = 0$；

(4) 凡Borel集关于任意 $\alpha(x)$ 都是L-S可测集.

有了L-S可测集和测度后，我们就可以在其上建立相当于L可测函数和L积分的L-S可测函数和L-S积分的概念与有关理论，只需要将L可测换成L-S可测，L测度换成L-S测度，L零测度集换成L-S零测度集即可，这样得到的积分叫做 $f(x)$ 关于 $\alpha(x)$ 的L-S积分，记为 $\int_E f(x)\mathrm{d}\alpha(x)$，其中 $E \subseteq \mathbf{E}^1$ 为关于 $\alpha(x)$ 的L-S可测集，$f(x)$ 为 E 上的L-S可积或L-S积分确定的函数.

显然，当 $\alpha(x) = x$ 时，L-S可测函数与L-S积分就分别成为L可测函数与L积分了.

注意 以上所述L-S测度与L-S积分都是在 $\mathbf{E}^1$ 中的，可以推广到 $\mathbf{E}^n (n > 1)$ 中去.

习 题 1-5

1. 设 P_0 为[0, 1]上的Cantor集，$f(x) = \begin{cases} 1, & \text{当 } x \in P_0 \text{ 时}; \\ 2, & \text{当 } x \in ([0, 1] - P_0) \text{ 时}, \end{cases}$ 计算 $\int_0^1 f(x)\mathrm{d}x$.

2. 设函数 $f(x)$ 在[0, 1]上的Cantor集 P_0 中的点处等于0，而在 P_0 的具有长度为 $\frac{1}{3^n}$ 的余区间上等于 n $(n = 1, 2, \cdots)$，试证 $f(x)$ 在[0, 1]上可积，并计算 $\int_0^1 f(x)\mathrm{d}x$.

3. 设 E 是[a, b]的可测子集，则 E 的特征函数 $\chi_E(x)$ 在[a, b]上可积，且 $\int_a^b \chi_E(x)\mathrm{d}x = mE$.

4. 设 $f(x)$ 是 E 上的非负可测函数，又 E 中使 $f(x) \geqslant c$ 的那些点所成之集 A 的测度 $mA = a$，证明

$$\int_E f(x)\mathrm{d}x \geqslant ac.$$

5. $f(x)$ 是$[a, b]$上非负连续函数，如果$\int_a^b f(x)\mathrm{d}x = 0$，则 $f(x) \equiv 0$.

6. 对于 Lebesgue 意义下的上、下积分而言，相应于 Darboux 定理的结论是否成立？

7. 设 $mE < +\infty$，$f(x)$ 是 E 上可测函数，$E_n = E[n-1 \leqslant f < n]$，则 $f(x)$ 在 E 上可积的充分必要条件是 $\sum_{n=-\infty}^{+\infty} |n| mE_n < +\infty$.

8. 设$\{f_n(x)\}$是 E 上非负可积函数列，若 $\int_E f_n(x)\mathrm{d}x \to 0 (n \to \infty)$，则 $f_n(x) \Rightarrow 0$ 于 E.

9. 设 $mE < +\infty$，$\{f_n(x)\}$ 是 a.e. 有限可测函数列，证明：

$$\int_E \frac{|f_n(x)|}{1+|f_n(x)|}\mathrm{d}x \to 0 (n \to \infty)$$

的充分必要条件是 $f_n(x) \Rightarrow 0$ 于 E.

10. 设 $f(x) = \left(\frac{1}{x}\right)^{\alpha} \sin\frac{1}{x}$，$0 < x \leqslant 1$，讨论 α 为何值时，$f(x)$为$(0, 1]$上 L 可积函数或不可积函数.

11. 设由$[0, 1]$中取出 n 个可测子集 $E_1, E_2, \cdots, E_n$，假定$[0, 1]$中任一点至少属于这 n 个集中的 q 个，试证必有一集，它的测度大于或等于$\frac{q}{n}$.

12. 设 $mE \neq 0$，$f(x)$在 E 上可积，如果对于任何有界可测函数 $\varphi(x)$，都有

$$\int_E f(x)\varphi(x)\mathrm{d}x = 0,$$

则 $f(x) = 0$ a.e. 于 E.

13. 设 $f(x)$ 为$[0, 1]$上的可积函数，如果对于任意 $c(0 \leqslant c \leqslant 1)$，总有$\int_0^c f(x)\mathrm{d}x = 0$，则 $f(x) = 0$ a.e. 于$[0, 1]$.

14. 证明：$\lim\limits_{n\to\infty}\int_0^{+\infty}\left(1+\frac{t}{n}\right)^{-n} t^{-\frac{1}{n}}\mathrm{d}t = 1.$

15. 试从 $\frac{1}{1+x} = (1-x)+(x^2-x^3)+\cdots$，$0 < x < 1$，证明

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots.$$

16. 证明：

$$\int_0^1 \frac{x^p}{1-x}\ln\frac{1}{x}\mathrm{d}x = \sum_{n=1}^{+\infty}\frac{1}{(p+n)^2} (p > -1).$$

17. 设 $f_n(x)$为 E 上可积函数，$f_n(x) \xrightarrow{\text{a.e.}} f(x)$于 E，且

$$\int_E |f_n(x)|\mathrm{d}x < K \qquad (K \text{是常数}),$$

则 $f(x)$可积.

18. 设 $f(x)$在 $[a-\varepsilon, b+\varepsilon](\varepsilon > 0)$ 上可积，则

$$\lim_{t\to 0}\int_a^b |f(x+t) - f(t)|\mathrm{d}x = 0.$$

19. 设 $f(x)$在$(0, +\infty)$上可积，且 $f(x)$一致连续，则

$$\lim_{x\to+\infty} f(x) = 0.$$

20. 在 $S = \{(x, y) \mid -1 \leqslant x \leqslant 1, -1 \leqslant y \leqslant 1\}$上定义

$$f(x, y) = \begin{cases} \dfrac{xy}{(x^2+y^2)^2}, & \text{当 } x^2+y^2 \neq 0 \text{ 时；} \\ 0, & \text{当 } x^2+y^2 = 0 \text{ 时，} \end{cases}$$

则 $f(x, y)$的两个累次积分都存在且相等，但 $f(x, y)$在 S 上不可积.

21. 设 $f(x)$在$[a, b]$上绝对连续，且 $f'(x) \geqslant 0$ a. e. 于$[a, b]$，则 $f(x)$是单调递增函数.

22. 设连续函数 $f(x)$在$[a, b]$上除有限个点外，具有导数 $f'(x)$，且 $|f'(x)| \leqslant M$. 试证 $f(x)$是有界变差函数.

参 考 文 献

1. 程其襄，等. 实变函数与泛函分析基础. 北京：高等教育出版社，1983

2. 薛昌兴. 实变函数与泛函分析(上). 北京：高等教育出版社，1990

3. 江泽坚. 实变函数论. 北京：人民教育出版社，1961

4. 夏道行等. 实变函数论与泛函分析(上). 北京：人民教育出版社，1978

5. И. П. 那汤松. 实变函数论(上). 徐瑞云，译. 北京：人民教育出版社，1963

6. 张喜堂. 实变函数论的典型问题与方法. 武汉：华中师范大学出版社，2000

7. 魏勇. 让学生在学习《实变函数》过程中体会数学的创新方法. 数学教育学报，Vol. 9. No. 2，2000

8. 李文林. 数学史教程. 北京：高等教育出版社，施普林格出版社，2000

第二篇　一般拓扑学

拓扑学是一种新的几何学，它研究图形在连续变形下的不变的整体性质. 它是现代数学的一门重要基础课程. 在 20 世纪 30 年代开始形成. 现在，拓朴学已经形成了一般拓扑学(或称点集拓扑)、代数拓扑、微分拓扑等分枝. 它已经渗入到并沟通了数学的很多分枝，在自然科学和工程技术中有日益重要的作用.

本篇介绍一般拓扑学的基础知识，这是现代数学的重要基础知识.

第一章　度量空间

在高等代数和实变函数课程中，我们已经熟悉了 n 维欧氏空间. 度量空间是 n 维欧氏空间的十分自然的推广，是学习泛函分析、拓扑学这两门课程所必需的基础. 拓扑学要讨论的拓扑空间便是度量空间的广泛的、抽象的、自然的推广. 本章对度量空间进行较详细的介绍.

第一节　度量空间

一、度量空间的概念

极限是数学分析的基本概念之一. 各种不同形式的极限概念有不同形式的描述方式，在 n 维欧氏空间中，极限概念的描述必须用到距离的概念. 例如，所谓序列 $\{x_n\}$ 的极限是 x_0，便是指对于任意给定的正数 $\varepsilon > 0$，存在正整数 N，使得当 $n > N$ 时，所有的点 x_n 与 x_0 的距离小于 ε.

不少极限过程都能用距离描述，而且在论述极限的性质时，实质上也只是利用了距离的基本性质. 因此，为了进一步深入和研究各种极限过程，有必要抓住距离的基本性质，将距离这一概念抽象到一般的点集中.

定义 1　**设 S 是非空集合. $\rho: S\times S\to \mathbf{R}^1$ 是 S 上的二元实函数，满足以下的三条度量公理：**

(1) **正则性：** $\rho(x, y) \geqslant 0$，且 $\rho(x, y) = 0 \Leftrightarrow x = y$；

(2) **对称性：** $\rho(y, x) = \rho(x, y)$；

(3) **三角不等式：** $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$.

则 $\rho(x, y)$ 叫做 S 的点 x，y 的距离. 点集 S 按照距离 $\rho(x, y)$ 成为度量空间或距离空间，记为 (S, ρ). 函数 ρ 叫做度量空间 (S, ρ) 的距离函数或度量.

在不致混淆时，常将度量空间 (S, ρ) 简单地记作 S.

二、度量空间的例

度量空间包含十分广泛的数学概念. 下面罗列度量空间的一些重要例子，我们应当对它们十分熟悉.

例 1 设 $\mathbf{E}^n$ 是 n 维欧氏空间，则 $(\mathbf{R}^n, d)$ 是度量空间.

证 d 显然满足度量公理 1°、2°.

下证 d 满足度量公理 3°.

令 $u_i = x_i - z_i$，$v_i = x_i - y_i$，$w_i = y_i - z_i$，则 $u_i = v_i + w_i$.
故只须证

$$\sqrt{\sum_{i=1}^{n}(v_i + w_i)^2} \leqslant \sqrt{\sum_{i=1}^{n} v_i^2} + \sqrt{\sum_{i=1}^{n} w_i^2}. \tag{1}$$

不等式(1)叫做 Minkowski(**闵可夫斯基**)**不等式**.

不等式(1)与不等式

$$\sum_{i=1}^{n}(v_i + w_i)^2 \leqslant \sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 + 2\sqrt{\sum_{i=1}^{n} v_i^2} \cdot \sqrt{\sum_{i=1}^{n} w_i^2}$$

显然等价，因而就是与不等式

$$\sum_{i=1}^{n} v_i w_i \leqslant \sqrt{\sum_{i=1}^{n} v_i^2} \cdot \sqrt{\sum_{i=1}^{n} w_i^2} \tag{2}$$

等价.

当 $\sum_{i=1}^{n} v_i w_i \leqslant 0$ 时，不等式(2)显然成立，从而 Minkowski 不等式成立. 当 $\sum_{i=1}^{n} v_i w_i > 0$ 时，不等式(2)与不等式

$$\left(\sum_{i=1}^{n} v_i w_i\right)^2 \leqslant \sum_{i=1}^{n} v_i^2 \cdot \sum_{i=1}^{n} w_i^2 \tag{3}$$

等价. 不等式(2)与(3)都叫做 Cauchy-Schwarz(**柯西-施瓦兹**)**不等式**. 剩下的是去证明不等式(3)成立.

记 $$A = \sum_{i=1}^{n} v_i^2,\ B = \sum_{i=1}^{n} v_i w_i,\ C = \sum_{i=1}^{n} w_i^2$$

那么,含 x, y 的二次型

$$Ax^2 + 2Bxy + Cy^2 = \sum_{i=1}^{n} (v_i x + w_i y)^2$$

非负,因此,系数 A, B, C 必满足

$$AC - B^2 \geqslant 0.$$

这便证得了不等式(3).

例 2 设 n 维欧氏空间 $\mathbf{E}^n$ 的点的集合 $\mathbf{R}^n$,在 $\mathbf{R}^n$ 上定义非负实函数

$$d_1(x, y) = \sum_{i=1}^{n} | x_i - y_i |$$

或

$$d_2(x, y) = \max_{1\leqslant i\leqslant n} | x_i - y_i |$$

则可验证($\mathbf{R}^n$, d_1),($\mathbf{R}^n$, d_2)都是度量空间.证明留给读者.对于 $n = 2, 3$,建议读者讨论 $d_1(x, y)$, $d_2(x, y)$的几何意义.

例 3 **空间 l^2.**

令 $l^2 = \left\{ x = \{x_n\} \in \mathbf{R}^\infty \,\middle|\, \sum_{n=1}^{\infty} x_n^2 \text{ 收敛} \right\}$

对于任意正整数 n,由 Minkowski 不等式

$$\sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \leqslant \sqrt{\sum_{i=1}^{n} x_i^2} + \sqrt{\sum_{i=1}^{n} (-y_i)^2}$$

可以看到,当 $x = \{x_n\} \in l^2$, $y = \{y_n\} \in l^2$ 时,必有$\{x_n - y_n\} \in l^2$,即级数 $\sum_{n=1}^{\infty} (x_n - y_n)^2$ 收敛. 令

$$d(x, y) = \sqrt{\sum_{n=1}^{\infty} (x_n - y_n)^2}$$

则显然 d 满足度量公理 1°、2°.

令 $z = \{z_n\} \in l^2$. 对于任意的正整数 n,由 Minkowski 不等式,有

$$\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} = \sqrt{\sum_{i=1}^{n} [(x_i - y_i) + (y_i - z_i)]^2}$$

$$\leqslant \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} + \sqrt{\sum_{i=1}^{n} (y_i - z_i)^2}.$$

令 $n\to\infty$，则得

$$d(x,\ z)\leqslant d(x,\ y)+d(y,\ z).$$

因此，d 满足度量公理 3°. 这样，$(l^2,\ d)$是一个度量空间.

例 4 连续函数空间$C[a,\ b]$.

设 $C[a,\ b]$是区间$[a,\ b]$上连续函数全体的集合. $\forall\, x,\ y\in C[a,\ b]$，记

$$\rho(x,\ y)=\max_{t\in[a,\ b]}\mid x(t)-y(t)\mid,$$

则易验证$(C[a,\ b],\ \rho)$是一个度量空间. 在这个度量空间中，$\rho(x,\ y)$表示 $C[a,\ b]$中两点 $x,\ y$ 的距离. 它表示平面上连续曲线 $\xi=x(t)$ 与 $\xi=y(t)$ 上横坐标相同的两点之间的最大距离(见图 2.1.1).

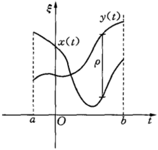


图 2.1.1

例 5 函数空间 $C^{(k)}[a,\ b]$.

设 k 是一个正整数，$C^{(k)}[a,\ b]$是区间$[a,\ b]$上具有连续的 k 阶导函数的函数全体的集合. $\forall\, x,\ y\in C^{(k)}[a,\ b]$，记

$$\rho_k(x,\ y)=\max_{0\leqslant j\leqslant k}\max_{a\leqslant t\leqslant b}\mid x^{(j)}(t)-y^{(j)}(t)\mid$$

其中，$x^{(0)}(t)$，$y^{(0)}(t)$ 分别表示$x(t)$，$y(t)$. 则易验证$(C^{(k)}[a,\ b],\ \rho_k)$是一个度量空间.

我们还可以把空间 $C[a,\ b]$视作空间 $C^{(0)}[a,\ b]$，这样例 5 中的 k 便可以看作非负整数亦可.

例 6 设 S 是非空集合. $\forall\, x,\ y\in S$，定义

$$\rho(x,\ y)=\begin{cases}0, & \text{当 } x=y \text{ 时;}\\ 1, & \text{当 } x\neq y \text{ 时,}\end{cases}$$

则$(S,\ \rho)$是一个度量空间，叫做离散度量空间.

从以上的例子可以看出，同一个集合，可以引入不同的距离的概念，从而得到不同的度量空间.

例 7 设$(S,\ \rho)$是一个度量空间. 定义

$$\rho_1(x,\ y)=\frac{\rho(x,\ y)}{1+\rho(x,\ y)}$$

或

$$\rho_2(x,\ y)=\begin{cases}\rho(x,\ y), & \text{当 } \rho(x,\ y)\leqslant 1 \text{ 时;}\\ 1, & \text{当 } \rho(x,\ y)>1 \text{ 时,}\end{cases}$$

则可以验证(S, ρ_1)，(S, ρ_2)都是度量空间.

例 8 设$C[a, b]$是区间$[a, b]$上连续函数全体的集合. $\forall f, g \in C[a, b]$，记

$$\rho(f, g) = \int_a^b |f(x) - g(x)| \, \mathrm{d}x.$$

显然，$\rho(f, g) \geqslant 0$，且当$f = g$时，$\rho(f, g) = 0$.

反之，假定$f \neq g$，则$\exists x_0 \in [a, b]$，使得$|f(x_0) - g(x_0)| > 0$. 记$|f(x_0) - g(x_0)| = A$，由$|f(x) - g(x)|$的连续性知，存在区间$I \subseteq [a, b]$，$x_0 \in I$，使得$\forall x \in I$有$|f(x) - g(x)| > \frac{A}{2}$，从而

$$\begin{aligned}\rho(f, g) &= \int_a^b |f(x) - g(x)| \, \mathrm{d}x \\ &\geqslant \int_I |f(x) - g(x)| \, \mathrm{d}x \geqslant \frac{A}{2} \cdot |I| > 0\end{aligned}$$

(其中$|I|$表示区间I的长度).

因此，当$\rho(f, g) = 0$时，只能有$f = g$.

再由

$$|f(x) - h(x)| \leqslant |f(x) - g(x)| + |g(x) - h(x)|$$

可推出

$$\int_a^b |f(x) - h(x)| \, \mathrm{d}x \leqslant \int_a^b |f(x) - g(x)| \, \mathrm{d}x + \int_a^b |g(x) - h(x)| \, \mathrm{d}x,$$

即$\rho(f, h) \leqslant \rho(f, g) + \rho(g, h)$

这样便论证了$(C[a, b], \rho)$是一个度量空间.

三、度量空间的子空间与积空间

设(S, ρ)是一个度量空间. A是S的一个非空子集，A的元素按照距离ρ(限制在$A \times A$上)，显然使(A, ρ)成一个度量空间，叫做(S, ρ)的子空间.

设(S_1, ρ_1)与(S_2, ρ_2)是两个度量空间. 记

$$S = S_1 \times S_2 = \{(x_1, x_2) \mid x_1 \in S_1, x_2 \in S_2\},$$

$$\rho((x_1, x_2), (y_1, y_2)) = \sqrt{\rho_1^2(x_1, y_1) + \rho_2^2(x_2, y_2)},$$

则易验证(S, ρ)是一个度量空间，叫做(S_1, ρ_1)与(S_2, ρ_2)的积空间，记为$(S_1, \rho_1) \times (S_2, \rho_2)$.

第二节　度量空间的点集

本节将讨论度量空间的一些特殊的点集. 因为度量空间是 n 维欧氏空间的自然推广，所以，我们从 n 维欧氏空间的点集理论出发，将邻域、开集、闭集等特殊的重要点集，以及序列的极限，点集的内点，聚点等重要概念引进到度量空间.

一、邻域

设 a 是度量空间(S, ρ)的一个点. δ 是一个正数，集合

$$O(a, \delta) = \{x \mid x \in S, \rho(x, a) < \delta\}$$

叫做点 a 在度量空间 S 的 δ-邻域，简称 a 的δ-邻域，或者 a 的邻域. a 叫做这个邻域的中心，δ 叫做这个邻域的半径.

集合 $\hat{O}(a, \delta) = O(a, \delta) - \{a\}$ 叫做点 a 的去心邻域.

定理 1　**设度量空间** (S, ρ), $x, y, z \in S$, **假设**

$$x \in O(y, \delta_1), x \in O(z, \delta_2),$$

则存在 $\delta > 0$, **使得**

$$O(x, \delta) \subseteq O(y, \delta_1) \cap O(z, \delta_2).$$

证　记 $d_1 = \delta_1 - \rho(x, y) > 0$, $d_2 = \delta_2 - \rho(x, z) > 0$, $\delta = \min\left\{\frac{d_1}{2}, \frac{d_2}{2}\right\}$, $\forall b \in O(x, \delta)$, 由

$$\begin{aligned}\rho(b, y) &\leqslant \rho(b, x) + \rho(x, y) \\ &< \delta + \rho(x, y) \\ &< d_1 + \rho(x, y) < \delta_1\end{aligned}$$

知　$b \in O(y, \delta_1)$.

同理可知 $b \in O(z, \delta_2)$.

故有定理 1 成立.

在度量空间(S, ρ)中，如果 $\exists x_0 \in S$, $\exists r > 0$，使得 $\forall x \in S$，有 $\rho(x, x_0) < r$，也就是 $x \in O(x_0, r)$，就说度量空间(S, ρ) 有界，否则，说度量空间(S, ρ) 无界.

二、序列的极限

定义 1　**设**$\{x_n\}$**是度量空间**(S, ρ)**的一个序列**，$x_0 \in S$. **如果** $\forall \varepsilon > 0$, $\exists$ **正**

整数 N,使得当 $n > N$ 时,有

$$\rho(x_n, x_0) < \varepsilon,$$

则说 x_0 是序列 $\{x_n\}$ 的极限. 又说序列 $\{x_n\}$ 按照距离 $\rho(x, y)$ 收敛于 x_0. 记为

$$\lim_{n\to\infty} x_n = x_0,$$

或

$$x_n \to x_0 (n \to \infty).$$

这时,称 $\{x_n\}$ 为收敛序列.

定理 2 在度量空间中,收敛序列的极限是惟一的.

证 假设在度量空间 (S, ρ) 中,x_0, y_0 都是收敛序列 $\{x_n\}$ 的极限,则

$$0 \leqslant \rho(x_0, y_0) \leqslant \rho(x_n, x_0) + \rho(x_n, y_0),$$

令 $n\to\infty$,有 $\rho(x_n, x_0)\to 0$, $\rho(x_n, y_0)\to 0$,故必有 $\rho(x, y) = 0$,因而 $x = y$.

定理 3 设 $\{x_n\}$, $\{y_n\}$ 是度量空间 (S, ρ) 的分别收敛于 x_0, y_0 的两个序列,那么,

$$\rho(x_n, y_n) \to \rho(x_0, y_0)\ (n \to \infty).$$

证 因为 $\rho(x_n, y_n) \leqslant \rho(x_n, x_0) + \rho(x_0, y_0) + \rho(y_0, y_n)$

以及

$$\rho(x_0, y_0) \leqslant \rho(x_0, x_n) + \rho(x_n, y_n) + \rho(y_n, y_0),$$

故有

$$|\rho(x_n, y_n) - \rho(x_0, y_0)| \leqslant \rho(x_n, x_0) + \rho(y_n, y_0),$$

令 $n\to\infty$,便得所要证明的结论.

例 1 连续函数空间 $C[a, b]$ 中,序列 $\{x_n\}$ 按照距离

$$\rho(x, y) = \max_{t\in[a, b]} |x(t) - y(t)|$$

收敛于点 x 就是连续函数列 $\{x_n(t)\}$ 在 $[a, b]$ 上一致收敛于函数 $x(t)$.

三、开集与闭集

定义 2 设 E 是度量空间 (S, ρ) 的一个点集, $a \in S$. 如果存在邻域 $O(a, \delta) \subseteq E$, 则说 a 是点集 E 的一个内点. 点集 E 的全体内点的集合,叫做 E 的内部,记为 $\mathrm{Int}E$.

定义 3 设 E 是度量空间 (S, ρ) 的一个点集,E 叫做开集,如果 E 的每个点都是 E 的内点.

定理 4 设度量空间(S, ρ). 则下列命题为真：

(1) S 与空集$\varnothing$是开集；

(2) 有限个开集的交集是开集；

(3) 任意多个开集的并集是开集.

证 S 的点显然是S 的内点,空集$\varnothing$不含不是内点的点,故 S, $\varnothing$都是开集.

设A, B是开集, $\forall x \in A \cap B$,则 x 是A 的内点,也是 B 的内点. 因此,存在邻域 $O(x, \delta_1) \subseteq A$,邻域$O(x, \delta_2) \subseteq B$. 由本节定理 1,存在邻域$O(x, \delta) \subseteq A \cap B$,故 x 是$A \cap B$ 的内点,从而 $A \cap B$ 是开集.

任意多个开集的并集的任意一点总是其中某个开集的内点,从而也是这个并集的内点,因而这个并集也是开集.

定义 4 设 E 是度量空间(S, ρ)的一个点集，$a \in S$. 如果 $\forall \delta > 0$, $O(\hat{a}, \delta) \cap E \neq \varnothing$,则说 a 是点集E 的一个聚点.

点集 E 的全体聚点的集合,叫做 E 的导集,记为 E'; $E \cup E'$ 叫做E 的闭包,记为 $\overline{E}$.

如果 $\overline{E} = S$, 则说 E 是 S 的稠密子集.

定理 5 设 E 是度量空间(S, ρ)的一个点集，$a \in S$，则下列命题等价：

(1) a 是E 的聚点；

(2) $\forall \delta > 0$, $O(a, \delta) \cap E$ 是无穷点集；

(3) 存在 E 的序列$\{x_n\}$(当 $i \neq j$ 时,$x_i \neq x_j$),使得

$$\lim_{n \to \infty} x_n = a.$$

定理 5 的证明完全可以逐字逐句照搬欧氏空间的点集 E 的聚点的类似性质的证明,这里从略了.

由定理 5 的命题(3),聚点又叫做极限点.

定理 6 E 是度量空间(S, ρ)的点集,则 $(E')' \subseteq E'$.

证 任取 $a \in (E')'$, $\forall \delta > 0$, $\exists b \in E' \cap O(\hat{a}, \delta)$,由 b 是E 的聚点,取 $\delta_1 = \min\left\{\frac{\rho(a, b)}{2}, \frac{\delta - \rho(a, b)}{2}\right\}$, $\exists c \in E \cap O(\hat{b}, \delta_1)$.

由于

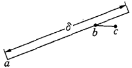


图 2.1.2

$$\rho(a, c) \leqslant \rho(a, b) + \rho(b, c)$$
$$< \rho(a, b) + (\delta - \rho(a, b)) = \delta,$$
$$\rho(a, c) \geqslant \rho(a, b) - \rho(b, c)$$
$$\geqslant \rho(a, b) - \frac{\rho(a, b)}{2} > 0,$$

故 $c \in E \cap O(\hat{a}, \delta)$，因此 a 是 E 的聚点，定理得证.

定理 7 设 A, B 是度量空间(S, ρ)的两个点集，则下列命题为真：

(1) $A \subseteq B \Rightarrow A' \subseteq B'$；

(2) $A' \cup B' = (A \cup B)'$.

证 (1) 任取 $a \in A'$. a 的 δ-邻域含有 A 的无穷多个点，从而也含有 B 的无穷多个点，故 $a \in B'$.

(2) 由 $A \subseteq A \cup B$，知 $A' \subseteq (A \cup B)'$，同理 $B' \subseteq (A \cup B)'$，从而

$$A' \cup B' \subseteq (A \cup B)'.$$

另一方面，任取 $c \in (A \cup B)'$，则 c 的 δ-邻域含有 $A \cup B$ 的无穷多个点，因而这个邻域含有 A 的或 B 的无穷多个点，即 $c \in A'$ 或 B'，从而

$$(A \cup B)' \subseteq A' \cup B'.$$

于是

$$A' \cup B' = (A \cup B)'.$$

定理 8 设 A, B 是度量空间(S, ρ)的两个点集，则下列命题为真：

(1) $\overline{\varnothing} = \varnothing$；

(2) $A \subseteq \bar{A}$；

(3) $\bar{\bar{A}} = \bar{A}$；

(4) $\overline{A \cup B} = \bar{A} \cup \bar{B}$.

证 (1) 显然 $\varnothing' = \varnothing$，从而 $\overline{\varnothing} = \varnothing \cup \varnothing' = \varnothing \cup \varnothing = \varnothing$.

(2) $A \subseteq A \cup A' = \bar{A}$.

(3)
$$\begin{aligned}\bar{\bar{A}} &= \bar{A} \cup (\bar{A})' = \bar{A} \cup (A \cup A')' \\ &= \bar{A} \cup (A' \cup (A')') \\ &= \bar{A} \cup A' = \bar{A}.\end{aligned}$$

(4)
$$\begin{aligned}\overline{A \cup B} &= (A \cup B) \cup (A \cup B)' \\ &= (A \cup B) \cup (A' \cup B') \\ &= (A \cup A') \cup (B \cup B') \\ &= \bar{A} \cup \bar{B}.\end{aligned}$$

从以上定理可以看出：**度量空间(S, ρ)的有限子集没有聚点. 如果 S 的无穷子集 E 的每相异两点的距离都大于一个固定的正数，则 E 没有聚点.**

定义 5 设 E 是度量空间(S, ρ)的一个点集，E 叫做闭集，如果 $E' \subseteq E$.

根据定义 5，闭集 E 的"闭"字的含意是指 E 的聚点被封闭在 E 中.

定理 9 设 E 是度量空间(S, ρ)的一个点集. 记 $E^c = S - E$，那么

$$E \text{ 是开集} \Leftrightarrow E^c \text{ 是闭集}.$$

证 若 E 是开集，则 $\forall a \in E$，a 是 E 的内点，因而 $\exists \delta > 0$，使得 $O(a, \delta) \subseteq$

E,即 $O(\hat{a}, \delta) \cap E^c = \varnothing$,故 $a \bar{\in} (E^c)'$,从而$(E^c)' \subseteq E^c$, E^c 为闭集.

反之,若 E^c 为闭集,则 $\forall a \in E$, $a \bar{\in} (E^c)'$,故 $\exists \delta > 0$,使得$O(a, \delta) \cap E^c = \varnothing$,即 $O(a, \delta) \subseteq E$, a 是 E 的内点,E 是开集.

由定理 9 及定理 4,立即可以推导出关于闭集的重要性质.

定理 10 **设度量空间(S, ρ). 则下列命题为真:**

(1) S **与空集**$\varnothing$**是闭集;**

(2) **有限个闭集的并集是闭集;**

(3) **任意多个闭集的交集是闭集.**

由上述论述还不难看出:度量空间(S, ρ)的任一点集 E 的**内部** $\text{Int}E$ **是开集**;E **的闭包** $\bar{E}$ **是闭集**,从而 E **是闭集的充分必要条件是** $E = \bar{E}$.

四、点集与点集的距离

设 A, B 是度量空间(S, ρ)的两个非空点集,我们将

$$\rho(A, B) = \inf\{\rho(x, y) \mid x \in A, y \in B\}$$

定义为点集 A 与点集 B 的距离. 我们还规定:**空集**$\varnothing$**与任一点集** A **的距离** $\rho(\varnothing, A)$ **为** 0. 我们将

$$\rho(x, B) = \inf\{\rho(x, y) \mid y \in B\}$$

定义为 S 的点 x 与点集 B 的距离.

我们还将

$$\sup\{\rho(x, y) \mid x \in A, y \in A\}$$

定义为点集 A 的直径.

定理 11 **设** E **是度量空间(S, ρ)的一个非空点集,** $x \bar{\in} \bar{E}$,**则** $\rho(x, E) > 0$.

证 因 $x \bar{\in} \bar{E}$,故 $x \in (\bar{E})^c$,而 $\bar{E}$ 是闭集,故$(\bar{E})^c$ 是开集. 因此, $\exists \delta > 0$,使得邻域 $O(x, \delta) \subseteq (\bar{E})^c$,从而 $\forall y \in E \subseteq \bar{E}$,均有 $\rho(x, y) \geqslant \delta$. 故 $\rho(x, E) = \inf\{\rho(x, y) \mid y \in E\} \geqslant \delta > 0$.

第三节 连续映射与拓扑映射

数学分析研究的对象是 n 维欧氏空间上的函数,主要是连续函数,我们已经从 n 维欧氏空间出发认识了度量空间. 也定义了度量空间的序列的极限. 现在,应该研究度量空间之间的、与连续函数相应的映射.

一、连续映射

定义 1 **设** X **与** Y **都是度量空间,映射** $f: X \to Y$, $x_0 \in X$. **如果** $\forall \varepsilon > 0$,

$\exists\delta>0$，使得当 $x\in O(x_0,\delta)$ 时，有 $f(x)\in O(f(x_0),\varepsilon)$，则说映射 f 在点 x_0 连续.

如果映射 f 在 X 的每一点处都连续，则说 f 在 X 上连续. 这时，映射 f 叫做是连续映射. X、Y 分别叫做映射 f 的定义空间与值空间. 特别地，当 $Y=\mathbf{E}^1$ 时，连续映射 f 通常叫做连续实值函数，或简称连续函数.

设 $A\subseteq X$，集合

$$f(A)=\{y\mid \exists x\in A,\text{使得 } y=f(x)\}$$

叫做 A 在 f 下的像.

设 $B\subseteq Y$，集合

$$f^{-1}(B)=\{x\mid \exists y\in B,\text{使得 } f(x)=y\}$$

叫做 B 在 f 下的原像.

定理 1 设 X 与 Y 都是度量空间. 映射 $f:X\to Y$. 则下列命题等价：

(1) f 是连续映射；

(2) Y 的开集在 f 下的原像是 X 的开集；

(3) Y 的闭集在 f 下的原像是 X 的闭集；

(4) $A\subseteq X\Rightarrow f(\bar{A})\subseteq\overline{f(A)}$；

(5) $\forall x_0\in X$，$\{x_n\}$ 是 X 中的任意收敛于 x_0 的序列，有

$$f(x_n)\to f(x_0)\ (n\to\infty).$$

证 (1)⇒(2)：$f:X\to Y$ 是连续映射，B 是 Y 的开集. 如果 $f^{-1}(B)$ 是空集，则 $f^{-1}(B)$ 是开集. 设 $f^{-1}(B)\neq\varnothing$，任取 $x_0\in f^{-1}(B)$，则 $f(x_0)\in B$. 由 B 是开集，$\exists\varepsilon>0$，使得邻域 $O(f(x_0),\varepsilon)\subseteq B$，又因 f 在 x_0 连续，故 $\exists\delta>0$，使得当 $x\in O(x_0,\delta)$ 时，$f(x)\in O(f(x_0),\varepsilon)\subseteq B$. 故 $O(x_0,\delta)\subseteq f^{-1}(B)$. 因而 $f^{-1}(B)$ 是开集.

(2)⇒(3)：设 B 是 Y 的闭集，则 $Y-B$ 是开集，因而 $f^{-1}(Y-B)$ 是开集. 但 $f^{-1}(Y-B)=X-f^{-1}(B)$. 故 $X-f^{-1}(B)$ 是开集，从而 $f^{-1}(B)$ 是闭集.

(3)⇒(4)：$\forall x_0\in\bar{A}$，如果 $x_0\in A$，显然 $f(x_0)\in f(A)\subseteq\overline{f(A)}$. 假设 $\exists x_0\in\bar{A}$，而 $f(x_0)\bar{\in}\overline{f(A)}$，则只可能 $x_0\in A'$，即 x_0 是 A 的聚点. 记 $B=\overline{f(A)}$，则 B 是闭集，从而 $f^{-1}(B)$ 是闭集，且因 $B=\overline{f(A)}\supseteq f(A)$，故 $A\subseteq f^{-1}(B)$. 故 x_0 是 A 的聚点也是 $f^{-1}(B)$ 的聚点. 由 $f^{-1}(B)$ 是闭集，应当有 $x_0\in f^{-1}(B)$. 因此 $f(x_0)\in B=\overline{f(A)}$. 矛盾.

(4)⇒(1)：假设 f 不是连续映射，不妨设 f 在 $x_0\in X$ 不连续，则 $\exists\varepsilon_0>0$，使得 $\forall\delta>0$，都存在点 $x_1\in O(\hat{x}_0,\delta)$，而 $f(x_1)\bar{\in} O(f(x_0),\varepsilon_0)$，即 $f(x_1)\in Y-O(f(x_0),\varepsilon_0)$，因而 x_0 是集合 $A=f^{-1}(Y-O(f(x_0),\varepsilon_0))$ 的聚点，故 $x_0\in$

$\bar{A}$. 注意到$Y-O(f(x_0), \varepsilon_0)$是闭集，$f(x_0) \bar{\in} Y-O(f(x_0), \varepsilon_0)$，而

$$f(A) = ff^{-1}(Y-O(f(x_0), \varepsilon_0)) \subseteq Y-O(f(x_0), \varepsilon_0).$$

即知 $x_0 \in \bar{A}$ 但 $f(x_0) \bar{\in} \overline{f(A)}$. 这与(4) 矛盾.

(1)⇒(5)：因 f 在点 x_0 连续，故 $\forall \varepsilon > 0$，$\exists \delta > 0$，使得当$x \in O(x_0, \delta)$时，$f(x) \in O(f(x_0), \varepsilon)$. 由$\lim\limits_{n\to\infty} x_n = x_0$ 知，$\exists$ 正整数 N，使得当 $n > N$ 时，$x_n \in O(x_0, \delta)$，因而$f(x_n) \in O(f(x_0), \varepsilon)$，这便是$\lim\limits_{n\to\infty} f(x_n) = f(x_0)$.

(5)⇒(1)：假设 f 不是连续映射，不妨设 f 在 $x_0 \in X$ 不连续. 则 $\exists \varepsilon_0 > 0$，对于任意收敛于 0 的正数数列$\{\delta_n\}$，总是 $\exists x_n \in O(x_0, \delta_n)$，使得 $f(x_n) \bar{\in} O(f(x_0), \varepsilon_0)$，这样便可选出一个收敛于 x_0 的序列$\{x_n\}$(当 $i \neq j$ 时，$x_i \neq x_j$)，而

$$\lim_{n\to\infty} f(x_n) \neq f(x_0),$$

这与(5)矛盾.

定理 1 从多方面刻画了连续映射，我们应当加深理解. 显然，度量空间的恒等映射是连续映射. 如果 $f: X \to Y$ 与 $g: Y \to Z$ 是度量空间之间的连续映射，则 $gf: X \to Z$ 也是连续映射，用连续映射的定义是不难验证的.

例 1 离散度量空间 X 到任一度量空间 Y 的映射 f 都是连续映射.

证 因离散度量空间的任意点集都没有聚点. 故其任意点集都是闭集，因而 Y 的闭集在 f 下的原像总是 X 的闭集. 故 f 是连续映射.

例 2 设度量空间(S, ρ)，则 $\rho: S \times S \to \mathbf{E}^1$ 是连续函数.

证 $\forall (x_0, y_0) \in S \times S$，任取$S \times S$的收敛于$(x_0, y_0)$的序列$\{x_n, y_n\}$，则 $x_n \to x_0$，$y_n \to y_0$. 由

$$\rho(x_n, y_n) \leqslant \rho(x_n, x_0) + \rho(x_0, y_0) + \rho(y_n, y_0)$$

及

$$\rho(x_0, y_0) \leqslant \rho(x_0, x_n) + \rho(x_n, y_n) + \rho(y_n, y_0),$$

有

$$|\rho(x_n, y_n) - \rho(x_0, y_0)| \leqslant \rho(x_n, x_0) + \rho(y_n, y_0),$$

令 $n \to \infty$，则得 $\rho(x_n, y_n) \to \rho(x_0, y_0)$.

故 ρ 是连续函数.

二、拓扑映射

定义 2 **设 X 与 Y 都是度量空间，映射 $f: X \to Y$ 叫做拓扑映射，如果 f 是双射，并且 f 及逆映射 f^{-1} 都是连续映射.**

如果两个度量空间 X 与 Y 之间存在拓扑映射 f，则说 X 与 Y 是同胚的，记为 $X \cong Y$.

易知，度量空间的同胚关系是一个等价关系.

如果度量空间的某性质，在每一拓扑映射下保持不变，则叫做拓扑性质.

由定理 1 可知：开集，闭集，子集的闭包与收敛序列都是拓扑性质.

例 3 将平面图形看作其点的集合，按照欧氏空间 $\mathbf{E}^2$ 的距离 d，平面图形都是度量空间($\mathbf{E}^2$, d)的子空间. 作为度量空间，圆周与椭圆、圆周与正方形的边界、圆的内部与 $\mathbf{E}^2$ 分别是同胚的.

例 4 设 $X = \{x \mid 0 \leqslant x < 2\pi\}$，$Y$ 是 $\mathbf{E}^2$ 中的单位圆周，以 $(1, \theta)$ 记单位圆周上的点的极坐标 $(0 \leqslant \theta < 2\pi)$，令 $f(x) = (1, x)$，显然 f 是 X 到 Y 的连续映射，且是双射. 但 f 却不是拓扑映射，因为 f^{-1} 不是连续映射，这从 X 的闭集 $A = [\pi, 2\pi)$ 的像 $f(A)$ 不是 Y 的闭集便知.

三、等距映射

定义 3 **设度量空间 (X, ρ) 与 (Y, d). 映射 $f: X \to Y$ 叫做等距映射，如果 f 是满射，$\forall a, b \in X$ 有**

$$d(f(a), f(b)) = \rho(a, b).$$

定理 2 **等距映射是拓扑映射.**

证 设 f 是度量空间 (X, ρ) 到 (Y, d) 的等距映射，先证 f 是单射. 这由

$$f(a) = f(b) \Rightarrow d(f(a), f(b)) = 0 \Rightarrow \rho(a, b) = 0 \Rightarrow a = b$$

便知. 故 f 是双射，有逆映射 f^{-1}.

再证 f^{-1} 是等距映射：

$\forall u, v \in Y$，记 $f^{-1}(u) = a$, $f^{-1}(v) = b$，则

$$\rho(f^{-1}(u), f^{-1}(v)) = \rho(a, b) = d(f(a), f(b)) = d(u, v),$$

由此便知 f^{-1} 是等距映射.

最后证等距映射是连续映射：

因为 $\forall \varepsilon > 0$, $\exists \delta = \varepsilon$，使得

$$\rho(a, b) < \delta \Rightarrow d(f(a), f(b)) = \rho(a, b) < \varepsilon.$$

这样便知等距映射 f 是连续映射. 由于 f^{-1} 也是连续映射. 因此，f 是拓扑映射.

如果度量空间的性质在每一个等距映射下保持不变，但却不能在每个拓扑映射下保持不变，则叫做度量性质.

第四节 列紧度量空间

在 n 维欧氏空间的点集理论中，聚点原理说的是“有界无穷点集必有聚点”. 这个聚点不一定属于这个点集，但一定属于这个欧氏空间. 与聚点原理等价的两

个命题是“有界序列必有收敛子列”以及有限覆盖定理. 有限覆盖定理说的是“有限闭区间的任一开覆盖一定有一个有限的子覆盖”. 这几个等价的命题在数学分析的基础理论中起着重大作用.

本节将在适当的度量空间建立与这些命题相应的理论.

一、列紧度量空间

定义 1　度量空间叫做列紧度量空间,如果它的每一个无穷子集都有聚点.

例 1　欧氏空间 $\mathbf{E}^1$ 上的闭区间$[a, b]$作为度量空间是列紧度量空间.

若 S 是非空有限集,则度量空间(S, ρ)是列紧度量空间.

作为度量空间,欧氏空间 $\mathbf{E}^1$ 以及 $\mathbf{E}^1$ 上的开区间(a, b),半开区间$[a, b)$, $(a, b]$都不是列紧度量空间.

定义 2　度量空间(S, ρ), S 的子集 A 叫做列紧子集,如果度量空间(A, ρ)是列紧度量空间.

由定义可知,度量空间(不必是列紧的),其列紧子集 A 的每个无穷子集都有聚点,并且这聚点一定属于 A. 例如,闭区间$[a, b]$是欧氏空间 $\mathbf{E}^1$ 的列紧子集,而开区间(a, b),半开区间$[a, b)$与$(a, b]$都不是 $\mathbf{E}^1$ 的列紧子集.

定理 1　度量空间(S, ρ)的列紧子集 A 是 S 的有界闭集.

证　先证 A 有界. 假设 A 无界,取点 $a_1 \in A$,则有 $a_2 \in A$ 使得$\rho(a_1, a_2) > 1$,有 $a_3 \in A$,使得 $\rho(\{a_1, a_2\}, a_3) > 1$, …,有 $a_{n+1} \in A$,使得 $\rho(\{a_1, a_2, \cdots, a_n\}, a_{n+1}) > 1$, …. 这样,$A$ 的无穷点集$\{a_1, a_2, \cdots, a_n, \cdots\}$无聚点. 这与 A 列紧矛盾.

再证 A 是闭集. 若 $A' = \varnothing$,则 A 是闭集. 若 $A' \neq \varnothing$, $\forall a \in A'$, a 是 A 的聚点. 那么,A 中有序列$\{x_n\}$, $i \neq j$ 时,$x_i \neq x_j$, $\{x_n\}$ 以 a 为极限,因此,a 也是 A 的无穷点集$\{x_1, x_2, \cdots, x_n, \cdots\}$的聚点. 由 A 是列紧子集便知 $a \in A$,从而 $A' \subseteq A$, A 是闭集.

定理 1 的逆命题不真. 例如,欧氏空间 $\mathbf{E}^1$ 上的半开区间 $S = (0, 1)$ 作为 $\mathbf{E}^1$ 的子空间,是一个度量空间. $A = S$ 是度量空间 S 的有界闭集. 但 A 不是 S 的列紧子集. 这是因为 A 的无穷子集$\left\{1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots\right\}$没有聚点.

定理 2　列紧度量空间(S, ρ)的子集 A 若是闭集,则 A 是 S 的列紧子集.

证　设 A 有无穷子集 B,则 B 也是 S 的无穷子集,由 S 列紧,B 有聚点 $b \in S$. 并且 $b \in \bar{B} \subseteq \bar{A} = A$, 即 A 是 S 的列紧子集.

定理 2 中的条件“S 是列紧度量空间”如果改为“S 是度量空间”,则结论一般不成立. 例如,n 维欧氏空间 $\mathbf{E}^n$ 中,$\mathbf{E}^1$ 是闭集,但不是 $\mathbf{E}^n$ 的列紧子集.

定理 3　n 维欧氏空间 $\mathbf{E}^n$ 的子集 A 是列紧子集$\Leftrightarrow$$A$ 是 $\mathbf{E}^n$ 的有界闭集.

证 必要性：由定理 1 即知.

充分性：若 A 有无穷子集 B，则 B 是 $\mathbf{E}^n$ 的有界无穷点集，故必有聚点. 因此，A 是列紧子集.

定理 4 **度量空间 (S, ρ) 是列紧度量空间的充分必要条件是 S 的任一序列都有收敛子列.**

证 必要性：S 的序列 $\{x_n\}$ 有两种情况. 第一种情况是序列 $\{x_n\}$ 的所有点的集合是无穷点集 A，则 A 有聚点 a，从而 $\{x_n\}$ 有子列 $\{x_{n_k}\}$ 收敛于 a. 第二种情况是序列的所有点的集合是有限集. 这样，至少有某一项的点在序列各项中重复出现无限多次. 因而 $\{x_n\}$ 有一各项相同的点构成的子列（叫做常项序列），当然是收敛子列.

充分性：设 A 是 S 的无穷子集，则在 A 中可找一子列 $\{x_n\}$，$i \neq j$ 时，$x_i \neq x_j$，而 $\{x_n\}$ 有收敛子列，其极限便是 A 的聚点. 故 S 是列紧度量空间.

定义 3 **设 ε 是一个正数，度量空间 (S, ρ) 的有限子集 W 叫做 S 的一个 ε-网，如果 $\forall x \in S$，有 $\rho(x, W) < \varepsilon$.**

定理 5 **对于任意正数 ε，列紧度量空间 (S, ρ) 有 ε-网.**

证 假设对某个正数 ε，S 无 ε-网. 任取 $a_1 \in S$，因 $\{a_1\}$ 不是 S 的 ε-网，故 $\exists a_2 \in S$，使得 $\rho(a_2, a_1) \geqslant \varepsilon$；又因 $\{a_1, a_2\}$ 不是 S 的 ε-网，故 $\exists a_3 \in S$，使得 $\rho(a_3, \{a_1, a_2\}) \geqslant \varepsilon$；…；如此下去，因 $\{a_1, a_2, \cdots, a_n\}$ 不是 S 的 ε-网，故 $\exists a_{n+1} \in S$，使得 $\rho(a_{n+1}, \{a_1, a_2, \cdots, a_n\}) \geqslant \varepsilon$；…. 这样，得到 S 的一个无穷子集

$$A = \{a_1, a_2, \cdots, a_n, \cdots\}$$

A 中任意两点 a_i，a_j 的距离 $\rho(a_i, a_j) \geqslant \varepsilon\ (i \neq j)$. 故 A 无聚点. 这与 (S, ρ) 是列紧度量空间矛盾.

应当指出，定理 5 的逆命题不真. 有 ε-网的度量空间不一定是列紧度量空间. 例如，$\mathbf{E}^1$ 上的半开区间 $(a, b]$ 有 ε-网，但却不是列紧的.

定理 6 **设 (S, ρ) 是列紧度量空间. 存在一个至多可数的 S 的点的邻域的集合 $\mathscr{M}$. $\forall$ 开集 $G \subseteq S$，$\exists \mathscr{M}_1 \subseteq \mathscr{M}$，使得 $G = \bigcup_{G_i \in \mathscr{M}_1} G_i$.**

证 对于每一个正整数 n，列紧度量空间 (S, ρ) 存在一个 $\frac{1}{n}$-网 $W_{\frac{1}{n}}$. 记 $W = \bigcup_{n=1}^{\infty} W_{\frac{1}{n}}$，则 W 是 S 的可数子集. 集合

$$\mathscr{M} = \{O(x, r) \mid x \in W,\ r \in \text{正有理数集 } \mathbf{Q}^+\}$$

是以 W 的点为中心，以有理数为半径的邻域的集合，是一个可数集.

$\forall$ 开集 $G \subseteq S$. 如果 $G = S$，则 $\mathscr{M}$ 有子集 $\mathscr{M}_1 = \mathscr{M}$ 使得 $G = \bigcup_{G_i \in \mathscr{M}_1} G_i$. 如果 $G \neq S$，$\forall x \in G$，由 G 是开集便知 $d = \rho(x, S-G) > 0$. 现在 W 中取一点 a，

使得 $\rho(x, a) < \frac{1}{3}d$. 那么,可取 $r \in \mathbf{Q}^+$, 满足 $\frac{1}{3}d < r < \frac{2}{3}d$. 则有

$$x \in O(a, r) \subseteq G.$$

定理 6 所述的列紧度量空间的性质叫做可数邻域族性质. 从定理的证明可以看出,定理 6 的条件可以适当削弱,事实上,有

定理 7　有一个可数的稠密子集的度量空间(叫做可分的度量空间),具有可数邻域族性质.

例 2　l^2 是可分的度量空间,具有可数的邻域族性质.

证　记

$$A = \{\{x_n\} \mid x_n \text{ 是有理数且至多有限项非零}\},$$

则 A 是 $\mathbf{E}^\omega$ 的可数子集. 设 $x = \{x_n\} \in l^2$, $\forall \varepsilon > 0$, $\exists$ 正整数 N, 使得

$$\sum_{n>N} x_n^2 < \varepsilon^2.$$

取 x_i' 满足 $|x_i - x_i'| < \frac{\varepsilon}{N}$ $(i = 1, 2, \cdots, N)$,则点 $x' = (x_1', x_2', \cdots, x_N', 0, 0, 0, \cdots) \in A$ 满足 $\rho(x, x') < 2\varepsilon$.

故 $\rho(x, A) = 0$, 因而 A 是 l^2 的一个可数的稠密子集.

二、紧致性

定义 4　设 A 是度量空间 (S, ρ) 的一个子集, 2^S 是 S 的幂集, $\mathscr{M} \subseteq 2^S$. 如果 $A \subseteq \bigcup_{B \in \mathscr{M}} B$, 则说 $\mathscr{M}$ 是 A 在 S 中的一个覆盖.

特别地,

当 $A = S$ 时,就说 $\mathscr{M}$ 是 A 的一个覆盖;

当 $\mathscr{M}$ 的元素都是开集时,就说 $\mathscr{M}$ 是开覆盖;

当 $\mathscr{M}$ 的元素都是闭集时,就说 $\mathscr{M}$ 是闭覆盖;

当 $\mathscr{M}$ 是可数集时,就说 $\mathscr{M}$ 是可数无穷的覆盖;

当 $\mathscr{M}$ 是有限集时,就说 $\mathscr{M}$ 是有限覆盖;

当 $\mathscr{M}$ 的子集 $\mathscr{M}_1$ 也是 A 的在 S 中的一个覆盖时,就说 $\mathscr{M}_1$ 是 $\mathscr{M}$ 的一个子覆盖.

定理 8　设列紧度量空间 (S, ρ) 有一个开覆盖 $\mathscr{M}$, 则 $\exists$ 正数 λ, 使得 $\forall A \subseteq S$, 当 A 的直径小于 λ 时, $\exists G \in \mathscr{M}$, 使得 $A \subseteq G$.

证　假设结论不成立,则对任一正整数 n, $\exists A_n$, A_n 的直径小于 $\frac{1}{n}$, 使得 $\forall G \in \mathscr{M}$ 都有 $A_n \nsubseteq G$. 在每个 A_n 任取一点 a_n. 则得 S 的一个序列 $\{a_n\}$. 由 S 是

列紧的，$\{a_n\}$有一个收敛子列. 设这个收敛子列的极限是 a. 因 $\mathscr{M}$ 是 S 的开覆盖，故 $\exists G \in \mathscr{M}$，使得 $a \in G$，由 G 是开集，有 $d = \rho(a, S-G) > 0$. 在上述$\{a_n\}$的收敛于 a 的收敛子列中可以选出一项 a_m，满足 $\rho(a, a_m) < \frac{d}{2}$，且 $m > \frac{2}{d}$. 现考察 A_m，由于 A_m 的直径小于$\frac{1}{m}$，所以对 $\forall x \in A_m$，有

$$\rho(a, x) \leqslant \rho(a, a_m) + \rho(a_m, x) < \frac{d}{2} + \frac{1}{m} < d.$$

故 $A_m \subseteq G$，矛盾.

定理 8 中的正数 λ 叫做 $\mathscr{M}$ 的一个Lebesgue 数.

有了上述准备，我们可以论证被叫做列紧度量空间的紧致性的性质.

定理 9　度量空间(S, ρ)是列紧度量空间$\Leftrightarrow S$ 的每一个开覆盖 $\mathscr{M}$ 有一个有限子覆盖.

证　必要性：设 λ 是 S 的开覆盖 $\mathscr{M}$ 的一个 Lebesgue 数. $A = \{a_1, a_2, \cdots, a_k\}$ 是 S 的一个$\frac{\lambda}{3}$-网，则 $S \subseteq \bigcup_{i=1}^{k} O\left(a_i, \frac{\lambda}{3}\right)$，由于$O\left(a_i, \frac{\lambda}{3}\right)$的直径$<\lambda$，故 $\exists G_i \in \mathscr{M}$，使得 $O\left(a_i, \frac{\lambda}{3}\right) \subseteq G_i$，$i = 1, 2, \cdots, k$. 因而 $\mathscr{M}_1 = \{G_1, G_2, \cdots, G_k\}$ 便是所求的有限子覆盖.

充分性：假设(S, ρ)不是列紧度量空间，则 S 有一个没有聚点的无穷点集 A. 因 $A' = \varnothing$，故 A 为闭集. 从而 $S-A$ 是开集. $\forall a \in A$，a 不是 A 的聚点，因而 $\exists \delta_a > 0$，使得 $O(a, \delta_a) \cap A = \{a\}$. 于是 $\mathscr{M} = \{S-A\} \cup \{O(a, \delta_a) \mid a \in A\}$ 是 S 的一个开覆盖. 显然，$\mathscr{M}$ 的真子集不能覆盖 S.

三、列紧度量空间上连续映射的性质

我们已经讨论了列紧度量空间，它可以说是 $\mathbf{E}^1$ 上的有限闭区间在度量空间上的推广. 如同研究数学分析中闭区间上连续函数的性质一样，现在我们讨论列紧度量空间上连续映射的性质.

定理 10　设 X 是列紧度量空间，Y 是度量空间，连续映射 $f: X \to Y$，则 $f(X)$是 Y 的列紧子集.

证　作为 Y 的子空间，$f(X)$是一个度量空间，$f: X \to f(X)$也是连续映射. 设 $\mathscr{U} = \{G_\alpha\}$ 是 $f(X)$ 的任一开覆盖. 则 $f^{-1}(G_\alpha)$ 是 X 的开集. $\mathscr{M} = \{f^{-1}(G_\alpha) \mid G_\alpha \in \mathscr{U}\}$ 是 X 的一个开覆盖. 由 X 列紧性，$\mathscr{M}$ 有一个有限的子覆盖

$$\mathscr{M}_1 = \{f^{-1}(G_{\alpha 1}), f^{-1}(G_{\alpha 2}), \cdots, f^{-1}(G_{\alpha n})\},$$

于是

$$\mathscr{U}_1 = \{G_{\alpha 1}, G_{\alpha 2}, \cdots, G_{\alpha n}\}$$

是 $f(X)$的一个有限覆盖,且是 $\mathscr{U}$ 的一个子覆盖.

定理 11　列紧度量空间上的连续函数有界,并且有最大值与最小值.

证　设 f 是列紧度量空间 X 上的连续函数. 即 $f: X \to \mathbf{E}^1$. 那么,$f(X)$是 $\mathbf{E}^1$ 上的有界闭集,显然包含了它的上确界与下确界.

定理 12　列紧度量空间 X 到度量空间 Y 的连续映射 f 是闭映射,即 X 的任一闭集 B 的象 $f(B)$是 Y 的闭集的映射.

证　B 是列紧度量空间 X 的闭子集,因而是列紧子集,于是 $f(B)$是 Y 的列紧子集,故 $f(B)$是 Y 的闭子集.

定理 13　列紧度量空间 X 到度量空间 Y 的连续映射 f 如果是双射,则 f 是拓扑映射.

证　只须证 f 的逆映射 f^{-1}是连续映射即可. 任取 X 的闭子集 B,则$f(B)$是 Y 的闭子集. 但 $f(B)$就是 A 在 f^{-1}下的原象. 因此,f^{-1}是连续映射.

定义 5　设度量空间(X, ρ)与度量空间(Y, d),映射 $f: X \to Y$ 叫做是在 X 上的一致连续映射,如果 $\forall \varepsilon > 0$, $\exists \delta > 0$,使得 $\forall x', x'' \in X$,当 $\rho(x', x'') < \delta$ 时,有 $d(f(x'), f(x'')) < \varepsilon$.

容易知道一致连续映射是连续映射,反之不然.

定理 14　设列紧度量空间(X, ρ)与度量空间(Y, d),映射 $f: X \to Y$ 是连续映射,则 f 是在 X 上一致连续映射.

证　f 是连续映射,则 $\forall \varepsilon > 0$, $\forall x \in X$, $\exists \delta_x > 0$ 使得当 $x' \in O(x, \delta_x)$ 时有 $d(f(x), f(x')) < \frac{\varepsilon}{2}$,由于 $\mathscr{M} = \{O(x, \delta_x) \mid x \in X\}$ 是列紧度量空间 X 的一个开覆盖,故有一个 Lebesgue 数 $\delta > 0$,当 $\rho(x', x'') < \delta$时,则 $\exists x \in X$,使得 $x', x'' \in O(x, \delta_x)$, 于是

$$\begin{aligned} d(f(x'), f(x'')) &\leqslant d(f(x'), f(x)) + d(f(x), f(x'')) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

第五节　完备度量空间

一、Cauchy 序列与完备空间

在数学分析中,判定数列是否收敛,Cauchy 判敛准则常常起着重要作用. 在度量空间中,判定序列是否收敛,是否有相应的法则呢? 下面的论述可以回答这

一问题.

定理 1　设 $\{x_n\}$ 是度量空间 (S, ρ) 的一个收敛序列，则 $\{x_n\}$ 满足下述的 Cauchy 条件：

$\forall \varepsilon > 0$，$\exists$ 正整数 N，使得当 $m, n > N$ 时，有

$$\rho(x_m, x_n) < \varepsilon.$$

证　设 $x_0 \in S$ 是序列 $\{x_n\}$ 的极限. $\forall \varepsilon > 0$，$\exists$ 正整数 N，使得当 $n > N$ 时，有 $\rho(x_n, x_0) < \frac{\varepsilon}{2}$. 那么，当 $m, n > N$ 时，有

$$\rho(x_m, x_n) \leqslant \rho(x_m, x_0) + \rho(x_0, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

虽然如此，定理 1 的逆命题却不真. 这就是说，在度量空间中，满足 Cauchy 条件的序列不一定是收敛序列. 例如，作为 $\mathbf{E}^1$ 的子空间 $(0, 1]$，令 $x_n = \frac{1}{n}$，度量空间 $(0, 1]$ 的序列 $\{x_n\}$ 显然满足 Cauchy 条件，但却没有极限.

定义 1　度量空间的满足 Cauchy 条件的序列叫做Cauchy 序列.

由上述可知，在度量空间中，收敛序列是 Cauchy 序列，但 Cauchy 序列不一定是收敛序列；而在 n 维欧氏空间 $\mathbf{E}^n$ 中，Cauchy 序列一定是收敛序列.

定义 2　度量空间 (S, ρ) 叫做完备度量空间，如果 (S, ρ) 中的每个 Cauchy 序列都是收敛序列.

完备度量空间常简称为完备空间.

度量空间 (S, ρ) 的子空间 (E, ρ) 若是完备空间，就说 E 是度量空间 (S, ρ) 的一个完备集.

我们规定：空集是度量空间的完备集.

由前述可知，n 维欧氏空间 $\mathbf{E}^n$ 是完备空间.

例 1　Hilbert 空间 (l^2, d) 是完备空间.

证　任取 l^2 的一个 Cauchy 序列

$$\{x_n\} = \{x_1^{(n)}, x_2^{(n)}, \cdots, x_k^{(n)}, \cdots\},$$

对于任意正整数 k，数列

$$\{x_k^{(n)}\} = x_k^{(1)}, x_k^{(2)}, x_k^{(3)}, \cdots, x_k^{(n)}, \cdots$$

是 $\mathbf{E}^1$ 的一个 Cauchy 序列，因此收敛于 $\mathbf{E}^1$ 的一个点 $x_k^{(0)}$. 用 x_0 表示数列

$$x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \cdots, x_k^{(0)}, \cdots$$

的极限. 下证 $x_0 \in l^2$.

$\forall \varepsilon > 0$，$\exists$ 正整数 N，使得当 $m, n > N$ 时，有

$$d(x_m, x_n) < \varepsilon,$$

因此,对于任意正整数 p,有

$$\sum_{k=1}^{p}(x_k^{(m)} - x_k^{(n)})^2 < \varepsilon^2,$$

固定 m,令 $n \to \infty$, 得

$$\sum_{k=1}^{p}(x_k^{(m)} - x_k^{(0)})^2 \leqslant \varepsilon^2,$$

再令 $p \to \infty$, 得

$$\sum_{k=1}^{\infty}(x_k^{(m)} - x_k^{(0)})^2 \leqslant \varepsilon^2,$$

于是,数列 $\{(x_k^{(m)} - x_k^{(0)})^2\} \in l^2$,即 $x_m - x_0 \in l^2$,因此 $x_0 \in l^2$.

当 $n > N$ 时,由

$$\begin{aligned} d(x_n, x_0) &\leqslant d(x_n, x_{(n+1)}) + d(x_{(n+1)}, x_0) \\ &< \varepsilon + \varepsilon = 2\varepsilon \end{aligned}$$

便知 l^2 的 Cauchy 序列$\{x_n\}$收敛于 x_0.

这样 l^2 便是完备空间.

二、完备度量空间的性质

定理 2 **设 E 是完备度量空间(S, ρ)的点集,那么,E 是 S 的完备集的充分必要条件是 E 是 S 的闭集.**

证 充分性:E 的任一 Cauchy 序列$\{x_n\}$也是完备度量空间(S, ρ)的 Cauchy 序列.因而,有极限 $x_0 \in S$,假若 $x_0 \overline{\in} E$,则 x_0 的任意邻域都含有 E 的异于 x_0 的点,因而 $x_0 \in E'$,这与 E 是 S 的闭集矛盾.

必要性:设 x_0 是 E 的聚点,则 x_0 是 E 的某个序列$\{x_n\}$的极限.$\{x_n\}$显然是 E 的 Cauchy 序列,因而是 E 的收敛序列,故 $x_0 \in E$,这样便知 E 是 S 的闭集.

定理 3 **完备度量空间(S_1, ρ_1)与完备度量空间(S_2, ρ_2)的积空间(S, ρ)是完备度量空间.**

证 这里 $S = S_1 \times S_2$,

$$\rho((x_1, x_2), (y_2, y_2)) = \sqrt{\rho_1^2(x_1, y_1) + \rho_2^2(x_2, y_2)},$$

任取 S 的一个 Cauchy 序列

$$(x_1^{(1)}, x_2^{(1)}), (x_1^{(2)}, x_2^{(2)}), \cdots, (x_1^{(n)}, x_2^{(n)}), \cdots$$

$\forall \varepsilon > 0$, $\exists$ 正整数 N,使得当 $m, n > N$ 时,有

$$\rho((x_1^{(m)}, x_2^{(m)}), (x_1^{(n)}, x_2^{(n)})) < \varepsilon,$$

故
$$\rho_1(x_1^{(m)}, x_1^{(n)}) \leqslant \rho((x_1^{(m)}, x_2^{(m)}), (x_1^{(n)}, x_2^{(n)})) < \varepsilon.$$

因此,序列

$$x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}, \cdots$$

是完备度量空间 S_1 的一个 Cauchy 序列,从而是 S_1 的收敛序列,记 $\lim\limits_{n\to\infty} x_1^{(n)} = x_1^{(0)}$.

同理,有 $\lim\limits_{n\to\infty} x_2^{(n)} = x_2^{(0)}$.

由此,易得 $\lim\limits_{n\to\infty}(x_1^{(n)}, x_2^{(n)}) = (x_1^{(0)}, x_2^{(0)})$.

因此,S 的 Cauchy 序列是收敛序列,S 是完备度量空间.

定理 4 (闭球套定理)设度量空间(S, ρ)是完备空间.

$$S_n = \{x \mid \rho(x_n, x) \leqslant \varepsilon_n\} \ (n = 1, 2, \cdots)$$

是 S 中的一套闭球,且

$$S_1 \supseteq S_2 \supseteq \cdots \supseteq S_n \supseteq \cdots$$

以及 $\varepsilon_n \to 0 (n \to \infty)$,那么,有惟一的 $\xi \in S$,使得 $\xi \in \bigcap\limits_{n=1}^{\infty} S_n$.

证 球心所组成的序列 $\{x_n\}$ 是 Cauchy 序列. 这是因为当 $m \geqslant n$ 时,由 $x_m \in S_m \subseteq S_n$,有

$$\rho(x_n, x_m) \leqslant \varepsilon_n.$$

因为 $\varepsilon_n \to 0$,故 $\forall \varepsilon > 0$,存在正整数 N,使得当 $n \geqslant N$ 时,有 $\varepsilon_n < \varepsilon$. 于是,当 $m, n \geqslant N$ 时,有

$$\rho(x_n, x_m) < \varepsilon,$$

所以 $\{x_n\}$ 是 Cauchy 序列. 由于 S 是完备空间,$\{x_n\}$ 便是收敛序列. 设 $\{x_n\}$ 收敛于 $\xi \in S$. 由 $\rho(x_n, x_m) \leqslant \varepsilon_n$,令 $m \to \infty$,可得

$$\rho(x_n, \xi) \leqslant \varepsilon_n,$$

即 $\xi \in S_n \quad (n = 1, 2, \cdots)$,因此 $\xi \in \bigcap\limits_{n=1}^{\infty} S_n$.

假若还有 $\eta \in \bigcap\limits_{n=1}^{\infty} S_n$,则

$$\rho(x_n, \eta) \leqslant \varepsilon.$$

令 $n \to \infty$,得 $\rho(\xi, \eta) = 0$,故 $\xi = \eta$,唯一性得证.

习 题 2-1

1. 证明：度量空间中的三条度量公理与以下的两条公理等价：

(i) **正则性**：$\rho(x, y) \geqslant 0$，且 $\rho(x, y) = 0 \Leftrightarrow x = y$；

(ii) **三角不等式**：$\rho(x, z) \leqslant \rho(x, y) + \rho(z, y)$.

2. 证明：本章第一节例 2 中的 d_1，d_2 是 $\mathbf{R}^n$ 的距离.

3. 证明：本章第一节例 7 中的 (S, ρ_1)，(S, ρ_2) 都是度量空间.

4. 设 S 为复数列 $\{x_n\}$ 的全体的集合. 令

$$\rho(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{|x_n - y_n|}{1 + |x_n - y_n|},$$

证明：(S, ρ) 是一个度量空间.

5. 设 $C^{\infty}[a, b]$ 是区间 $[a, b]$ 上无限次可微函数的集合. 令

$$\rho(f, g) = \sum_{n=0}^{\infty} \frac{1}{2^n} \max_{[a, b]} \frac{|f^{(n)}(t) - g^{(n)}(t)|}{1 + |f^{(n)}(t) - g^{(n)}(t)|},$$

证明：$(C^{\infty}[a, b], \rho)$ 是一个度量空间.

6. 设 S 是欧氏空间 $\mathbf{E}^3$ 中的任一球面. $\forall x, y \in S$，规定 x，y 的距离 $\rho(x, y)$ 是 x，y 两点的大圆上以 x，y 为端点的劣弧的弧长. 证明：$(\mathbf{E}^3, \rho)$ 是一个度量空间，ρ 不是 $\mathbf{E}^3$ 的 Euclid 距离 d，且

$$d(x, y) \leqslant \rho(x, y) \leqslant \frac{\pi}{2} d(x, y).$$

7. 设 x，y 是 Hilbert 空间 $\mathbf{E}^{\omega}$ 的两点. 证明：存在点 $z \in \mathbf{E}^{\omega}$，使得

$$\rho(x, z) = \rho(y, z) = \frac{1}{2} \rho(x, y).$$

8. 7 题中的点 z 叫做点 x，y 的**中点**，在一般的度量空间 S 中，是否一定存在两点 x，y 的中点？

9. 证明：度量空间 (S, ρ) 的点集 E 的内部 $\mathrm{Int}E$ 是 E 所包含的所有的开集的并集.

10. 证明：度量空间 (S, ρ) 的点集 E 的闭包 $\overline{E}$ 是所有包含 E 的闭集的交集.

11. 证明：在度量空间 $C^{\infty}[a, b]$（见习题 2-1 第 5 题）中，序列 $\{x_n\}$ 按距离收敛于 $x(t)$ 的充分必要条件是对每个非负整数 p，$\{x_n^{(p)}(t)\}$ 在 $[a, b]$ 上一致收敛于 $x^{(p)}(t)$.

12. 证明本章第二节定理 5.

13. 设 A，B 是度量空间 (S, ρ) 的点集. 证明：$(A \cap B)' \subseteq A' \cap B'$. 举例说明 $(A \cap B)' = A' \cap B'$ 一般不成立.

14. 设 E 是度量空间 (S, ρ) 的闭集. 证明：必有一列包含 E 的开集 A_1，A_2，…，使得 $\bigcap_{n=1}^{\infty} A_n = E$.

15. 设 E 是度量空间 (S, ρ) 的一个点集，a 是一个正数. 证明：点集

$$O(E, a) = \{x \mid \rho(x, E) < a\}$$

是开集，而点集

$$\{x \mid \rho(x, E) \leqslant a\}$$

是闭集.

16. 设 A_1，A_2 是度量空间(S, ρ)的两个点集，且 $\rho(A_1, A_2) > 0$. 证明：存在开集 O_1，O_2，使得

$$A_1 \subseteq O_1,\ A_2 \subseteq O_2,\ O_1 \cap O_2 = \varnothing.$$

17. $\mathbf{R}$ 是实数集，f 是离散度量空间$(\mathbf{R}^n, \rho)$到 n 维欧氏空间$(\mathbf{R}^n, d)$的恒等映射，证明：f 是连续的双射，但其逆映射 f^{-1} 不是连续映射.

18. 设 f 是度量空间 X 到 n 维欧氏空间 $\mathbf{E}^n$ 的映射，$x \in X$，记

$$f(x) = (f_1(x), f_2(x), \cdots, f_n(x))$$

证明：f 是连续映射$\Leftrightarrow$每个 f_i 是连续函数.

19. 设映射 $\varphi:\mathbf{E}^1 \to \mathbf{E}^1$，记 $G = \{(x, \varphi(x)) \mid x \in \mathbf{E}^1\} \subseteq \mathbf{E}^2$，又记映射 $f:G \to \mathbf{E}^1$ 为 $f(x, \rho(x)) = x$，证明：f 是拓扑映射$\Leftrightarrow\varphi$是连续函数.

20. 设度量空间(S, ρ)，$A \subseteq S$，$x \in S$. 证明：距离 $\rho(x, A)$在 S 上连续.

21. 证明：作为 n 维欧氏空间 $\mathbf{E}^n$ 的子空间，n 维实心球与 n 维正方形是同胚的，$\mathbf{E}^n$ 与减去一个点的 n 维球是同胚的.

22. 设度量空间(X, ρ)与度量空间(Y, d)，映射 $f:X \to Y$，$x_0 \in X$，但 $x_0 \overline{\in} X'$. 证明：f 在 x_0 连续.

23. 用长度 $\varepsilon > 0$ 的开区间覆盖区间$[0, 1]$，最少要多少个？

24. 如果$\{I_n\}$是覆盖$[0, 1]$的一列开区间，用$|I_n|$表示 I_n 的长度. 证明：

$$\sum_{n=1}^{\infty} |I_n| > 1.$$

25. 设 X 与 Y 都是欧氏空间$(\mathbf{E}^n, d)$的列紧子空间.

$$S = \{f \mid f:X \to Y,\ f \text{ 是连续映射}\}.$$

$\forall f, g \in S$，定义

$$\rho(f, g) = \sup_{x \in X} d(f(x), g(x)).$$

证明：(S, ρ)是一个度量空间.

26. 证明：有收敛子列的 Cauchy 序列是收敛序列.

27. 设(S, ρ)是完全度量空间. E_i 是 S 的开集且是 S 的稠密子集 $(i = 1, 2, \cdots)$. 证明：$\bigcap_{n=1}^{\infty} E_i$ 是 S 的稠密子集.

第二章 拓扑空间

度量空间的内容与应用虽然已经十分广泛，但是它却不能包括像射影空间等在实际应用中广泛存在的空间. 因此，有必要将度量空间的概念加以推广. 本章研讨的拓扑空间就是度量空间的推广. 在拓扑空间中，两点之间一般没有距离的意义. 在这样的前提下，我们要建立序列的收敛，函数的连续性等数学分析中相应的概念，利用这些概念，讨论一些拓扑空间的性质. 这就是本章要研究的内容.

第一节 拓扑空间的概念

一、拓扑与开集

从度量空间的论述中可以看到，极限是通过距离描述的，但实际上是可以通过特殊的开集——邻域来描述的. 基于这一点，可以抽象出拓扑空间的概念.

定义 1 **设 X 是一个非空集合，$\mathscr{T}\subseteq 2^X$. $\mathscr{T}$ 叫做 X 的一个拓扑，如果 $\mathscr{T}$ 满足下列的开集公理：**

1° $X\in\mathscr{T}$, $\varnothing\in\mathscr{T}$;

2° **若** $A, B\in\mathscr{T}$, **则** $A\cap B\in\mathscr{T}$;

3° **若** $\mathscr{T}_1\subseteq\mathscr{T}$, $\mathscr{T}_1\neq\varnothing$, **则** $\bigcup\limits_{A\in\mathscr{T}_1} A\in\mathscr{T}$.

X 与 X 上的拓扑 $\mathscr{T}$ 叫做一个拓扑空间，记为 $(X, \mathscr{T})$，有时也简记为 X. 拓扑 $\mathscr{T}$ 的元素叫做 X 的开集.

由定义可知，拓扑空间就是在其上定义了开集的集合. 拓扑空间 $(X, \mathscr{T})$ 的开集具有性质：

(1) **X 与 $\varnothing$ 是开集；**

(2) **有限多个开集的交集是开集；**

(3) **任意多个开集的并集是开集.**

当 X 是有限集时，拓扑空间 $(X, \mathscr{T})$ 叫做有限拓扑空间；

当 X 是无限集时，拓扑空间 $(X, \mathscr{T})$ 叫做无限拓扑空间.

拓扑空间广泛存在. 我们已经熟知的几何空间、n 维欧氏空间、度量空间都是特殊的拓扑空间.

例 1 (S, ρ) 是度量空间，$\mathscr{T}_\rho$ 是 S 的开集全体的集合，则 $\mathscr{T}_\rho$ 是 S 的一个拓

扑，$(S, \mathscr{T}_\rho)$是一个拓扑空间. $\mathscr{T}_\rho$ 叫做度量 ρ 诱导出的拓扑. 说度量空间(S, ρ)是一个拓扑空间，就是专指$(S, \mathscr{T}_\rho)$.

特别地，n 维欧氏空间 $\mathbf{E}^n$ 是拓扑空间，其拓扑是由欧氏空间的距离 d 诱导出的拓扑，叫做自然拓扑或欧氏拓扑.

例 2 任何非空集合 X 都有一个开集个数最少的拓扑 $\mathscr{T} = \{X, \varnothing\}$. $(X, \mathscr{T})$ 叫做平庸拓扑空间，简称平庸空间. 平庸空间只有两个开集：X 与空集$\varnothing$.

例 3 任何非空集合 X 都有一个开集个数最多的拓扑 2^X. $(X, 2^X)$叫做离散拓扑空间，简称离散空间. 离散空间中每个子集都是开集.

例 4 设集合 $X = \{0, 1\}$，$\mathscr{T} = \{X, \varnothing, \{0\}\}$. 可以验证$\mathscr{T}$是 X 上的一个拓扑. 拓扑空间$(X, \mathscr{T})$叫做W. Sierpinski(希尔宾斯基)空间.

二、拓扑基

在度量空间中，邻域是特殊的开集，每个开集都是一族邻域的并集. 这一事实使我们提出拓扑基的概念.

定义 2 **设拓扑空间 $(X, \mathscr{T})$，$\mathscr{B} \subseteq \mathscr{T}$，$\mathscr{B}$ 叫做拓扑 $\mathscr{T}$的基，如果 X 的每个开集都可表为 $\mathscr{B}$ 的(某些)元素的并集.**

拓扑 $\mathscr{T}$的基 $\mathscr{B}$ 又叫做拓扑空间$(X, \mathscr{T})$的拓扑基，或者 X 的拓扑基.

定理 1 **X 是非空集合，$\mathscr{B} \subseteq 2^X$. $\mathscr{B}$ 是 X 上一个拓扑 $\mathscr{T}$的基的充分必要条件是 $\mathscr{B}$ 满足：**

(1) $X = \bigcup\limits_{B \in \mathscr{B}} B$；

(2) $\forall B_\alpha, B_\beta \in \mathscr{B}$，若 $x \in B_\alpha \cap B_\beta$，则 $\exists B_\gamma \in \mathscr{B}$，使得 $x \in B_\gamma \subseteq B_\alpha \cap B_\beta$.

证 充分性：当 $\mathscr{B}$ 满足(1)、(2)，记

$$\mathscr{T} = \{A \mid \exists \mathscr{B}_A \subseteq \mathscr{B} \text{ 使得 } A = \bigcup_{B \in \mathscr{B}_A} B\}.$$

也就是说，用 $\mathscr{B}$ 的元素，即 X 的子集，构做一切可能的并集，所有的这样的并集便是 $\mathscr{T}$的所有的元素. 用拓扑的定义便可验证 $\mathscr{T}$是 X 的拓扑.

必要性：若 $\mathscr{B}$ 是某个拓扑的基，则因 X 是开集，故 X 是 $\mathscr{B}$ 的某些元素的并集，由此便可知 $\mathscr{B}$ 满足(1).

如果 $B_\alpha, B_\beta \in \mathscr{B}$，$x \in B_\alpha \cap B_\beta$，由于 $B_\alpha \cap B_\beta$ 是开集，故 $B_\alpha \cap B_\beta$ 是 $\mathscr{B}$ 的某些元素的并集，故 $\exists B_\gamma \in \mathscr{B}$，使得 $x \in B_\gamma \subseteq B_\alpha \cap B_\beta$.

定理 2 设 X 是非空集合，$\mathscr{B}$ 是 X 上的某个拓扑 $\mathscr{T}$的基，则 $\mathscr{T}$由 $\mathscr{B}$ 惟一确定.

证 定理 1 证明中所表示的 $\mathscr{T}$是 X 上的拓扑，$\mathscr{B}$ 是 $\mathscr{T}$的基，设 $\mathscr{B}$ 还是 X 上的拓扑 $\mathscr{T}_1$ 的基，$\forall A \in \mathscr{T}_1$，$A$ 是 $\mathscr{B}$ 的某些元素的并集，故 $A \in \mathscr{T}$，$\mathscr{T}_1 \subseteq \mathscr{T}$. 又因 $\mathscr{B} \subseteq \mathscr{T}_1$，故 $\forall A \in \mathscr{T}$，$A$ 必是 $\mathscr{B}$ 的某些元素的并集，从而也是 $\mathscr{T}_1$ 的某些元素的并

集，故 $A \in \mathscr{T}_1$，于是 $\mathscr{T} \subseteq \mathscr{T}_1$，故有 $\mathscr{T} = \mathscr{T}_1$.

由拓扑基 $\mathscr{B}$ 确定的拓扑 $\mathscr{T}$ 叫做由拓扑基 $\mathscr{B}$ 诱导出的拓扑.

例 5 设 $X = \{a, b, c\}$，集合

$$\mathscr{A} = \{X, \{a, b\}, \{a, c\}, \varnothing\}$$

不能成为 X 的拓扑基，这是因为

$$a \in \{a, b\} \cap \{a, c\},$$

但却不存在 $B \in \mathscr{A}$ 使得

$$a \in B \subseteq \{a, b\} \cap \{a, c\}.$$

集合 $\mathscr{B} = \{X, \{a, b\}, \{a, c\}, \{a\}, \varnothing\}$ 成为 X 的一个拓扑基，易验证，$\mathscr{B}$ 也是 X 的一个拓扑.

三、邻域

在拓扑空间中，包含点 x 的开集 U，叫做 x 的一个**邻域**. 在度量空间中，点 x 的 δ-邻域$O(x, \delta)$是指与点 x 距离小于δ 的点的集合，这样的邻域有中心 x 与半径 δ，感觉十分直观. 因此，我们将度量空间的点的 δ-邻域叫做**球形邻域**，以便区别于拓扑空间所述的点的邻域. 此前，在度量空间中，邻域一词是专指球形邻域，包含点 x 的开集并非都说是 x 的邻域，就是指并非都是球形邻域. 现在，当把度量空间视为拓扑空间时，包含点 x 的开集应当看作 x 的邻域了. 在一般拓扑空间中，因无距离的定义，自然就无邻域中心、邻域半径可言.

应当说明，在许多文献中，将包含点 x 的开集 U 叫做点 x 的开邻域，而将包含了点 x 的一个开邻域的任意子集都叫做点 x 的邻域. 这与我们上述的邻域定义在使用上没有本质差异. 无非是为了方便，我们所说的邻域就是专指后者的开邻域而已. 但读者应了解这一点，以便在读各种文献时，对邻域这一概念有正确理解.

拓扑空间$(X, \mathscr{T})$的任意一点都有邻域，例如，开集 X 就是 X 的每一个点的邻域. 平庸拓扑空间$(X, \mathscr{T})$的任意一点都有且只有一个邻域，X 是这些点的共同的邻域. 离散拓扑空间中，任意独点集$\{a\}$都是开集，因而$\{a\}$是点 a 的一个邻域. 在本节例 4 中所述的 W. Sierpinski 空间中，点 0 有两个邻域$\{0\}$与 X，而点 1 只有一个邻域 X. 顺便指出，虽然点 0 有邻域$\{0\}$不包含点 1，但点 1 的邻域总是包含了点 0. 这种现象在度量空间是不会发生的，因为度量空间(S, ρ)中的任意两个相异点 a, b，总有球形邻域 $O(a, r)$与 $O(b, r)$不相交，其中 $r = \frac{1}{2}\rho(a, b)$.

这里，我们初步看到拓扑空间与度量空间的差别.

有了邻域的概念，我们可以定义点集的内点的概念. 设 E 是拓扑空间的点集，点 x 说是 E 的一个内点，如果存在 x 的邻域 $U \subseteq E$. 点集 E 的全体内点的集合，叫做 E 的内部，记为 $\mathrm{Int}E$.

这些概念的定义，从字面上看，与度量空间中同名概念的定义是完全一致的. 因此，仿照度量空间中的论述，我们易得

定理 3　拓扑空间的点集 E 是开集的充分必要条件是 $E = \mathrm{Int}E$.

定义 3　设 x 是拓扑空间$(X, \mathscr{T})$的一个点，$\{W_\alpha\}$是 x 的一族邻域的集合. $\{W_\alpha\}$叫做 X 的在点 x 处的邻域基，如果对于点 x 的任意邻域 U，存在一个 W_α，使得 $W_\alpha \subseteq U$.

按照定义，拓扑空间的每一个点处都有邻域基，例如，该点的所有邻域的集合. 在度量空间中，以某点为中心，正有理数为半径的球形邻域全体的集合是该点处的邻域基.

在定义了拓扑空间的点的邻域后，完全可以理解拓扑空间的点集的邻域了. 设 E 是拓扑空间的一个点集，包含 E 的开集叫做点集 E 的邻域. 拓扑空间$(X, \mathscr{T})$的开集 X 是空间任一点集的邻域. 独点集$\{a\}$的邻域也是点 a 的邻域，反之亦然.

四、序列的极限

一般拓扑空间虽然没有距离的概念，但有点的邻域的概念，这样，也可以定义序列的极限了.

定义 4　设$\{x_n\}$是拓扑空间$(X, \mathscr{T})$的一个序列，$x_0 \in X$. 如果对于点 x_0 的任意邻域 U，都存在正整数 N，使得当 $n > N$ 时，有 $x_n \in U$，则说 x_0 是序列$\{x_n\}$的极限，或者说序列$\{x_n\}$收敛于 x_0. 记为

$$\lim_{n\to\infty} x_n = x_0 \quad \text{或} \quad x_n \to x_0.$$

有极限的序列叫做收敛序列，否则叫做发散序列.

如果我们把邻域看作度量空间的球形邻域，那么，定义 4 与度量空间、欧氏空间中序列极限的定义完全一致，几乎是逐字逐句照搬. 但是拓扑空间中的点的邻域与度量空间的球形邻域相比较，更具有抽象性、一般性、广泛性，因而拓扑空间中序列的极限可能有度量空间中序列的极限不具有的一些奇特现象.

例 6　设 $X = \{0, 1\}$，$\mathscr{T} = \{X, \varnothing, \{0\}\}$，则拓扑空间$(X, \mathscr{T})$是 W. Sierpinski 空间. 我们将各项都相等的序列叫做常项序列，设每个 $x_n = 0$，则已知点 0 是常项序列$\{x_n\}$的一个极限. 由于点 1 只有一个邻域 X，而 X 包含 0，因此，常项序列$\{x_n\}$的每一项都属于点 1 的任意的邻域，因而，点 1 也是序列$\{x_n\}$的极限. 由此看出，一般拓扑空间中，收敛序列的极限可能不唯一.

在 W. Sierpinski 空间中，设

$$x_n = \begin{cases} 1, & \text{当 } n \text{ 是奇数时；} \\ 0, & \text{当 } n \text{ 是偶数时.} \end{cases}$$

由极限定义易验证，序列$\{x_n\}$收敛于 1.

这对于度量空间、欧氏空间中序列的收敛性而言是非常令人惊讶的.

五、聚点与导集

仿照度量空间中点集的聚点与导集的定义，我们将这两个概念推广到拓扑空间.

定义 5 **设 E 是拓扑空间$(X, \mathscr{T})$的一个点集，$x_0 \in X$. x_0 叫做点集 E 的聚点，如果对于 x_0 的任意邻域 U，有 $E \cap (U \backslash x_0) \neq \varnothing$.**

点集 E 的全体聚点的集合，叫做点集 E 的导集，记为 E'.

按照定义，x_0 是 E 的聚点，则 x_0 的任意邻域包含 E 的异于 x_0 的点，反之亦然.

定义 5 也是逐字逐句照搬度量空间的点集的聚点与导集的定义. 但在拓扑空间中，聚点与导集也可能产生度量空间中不具有的一些奇特现象.

例 7 设集合 $X = \{a, b, c\}$，$\mathscr{T} = \{X, \varnothing, \{a, b\}, \{c\}\}$，在拓扑空间$(X, \mathscr{T})$中，我们考察点集 $E = \{a\}$ 有哪些聚点.

首先，点集 E 不包含异于 a 的点. 因此，a 不是 E 的聚点. 其次，点 c 有邻域$\{c\}$不包含 E 的点，故 c 也不是 E 的聚点. 再考察点 b，它有两个邻域，X 与$\{a, b\}$，它们都包含点 a，点 a 属于 E 而异于 b. 按聚点定义，点 b 是点集 E 的聚点. 按照导集的定义，E 的导集 $E' = \{b\}$. 同样的道理还可推知，$(E')' = \{b\}' = \{a\}$.

由例 7，我们可以知道如下一些事实：

(1) 在度量空间中，点集 E 的聚点 x 的任意邻域包含 E 的无穷多个点；

在一般拓扑空间中，点集 E 的聚点 x 的邻域可能只包含 E 的有限多个点.

(2) 在度量空间中，有限点集无聚点；

在一般拓扑空间中，有限点集可能有聚点.

(3) 在度量空间中，点集 E，$(E')' \subseteq E'$ 成立；

在一般拓扑空间中，可能有点集 E，使得 $(E')' \subseteq E'$ 不成立.

(4) 在度量空间中，若点 x_0 是点集 E 的聚点，则有 E 的序列 $\{x_n\}$($i \neq j$ 时，$x_i \neq x_j$) 收敛于 x_0；

在一般拓扑空间中，若点 x_0 是点集 E 的聚点，则有 E 的序列$\{x_n\}$收敛于 x_0(但序列$\{x_n\}$中的点可能不是两两互异).

定理 4 **设 A, B 是拓扑空间$(X, \mathscr{T})$的两个点集，则下列命题为真：**

(1) $\varnothing' = \varnothing$;

(2) $A \subseteq B \Rightarrow A' \subseteq B'$;

(3) $A' \cup B' = (A \cup B)'$;

(4) $(A')' \subseteq A' \cup A$.

证 (1)，(2)以及 $A' \cup B' = (A \cup B)'$ 都是显然的，下证

$$(A \cup B)' \subseteq A' \cup B'.$$

假设 $a \,\overline{\in}\, A' \cup B'$，则 $a \,\overline{\in}\, A'$ 且 $a \,\overline{\in}\, B'$. 由聚点定义，存在点 a 的邻域 U_1，U_2 使得

$$(U_1 - \{a\}) \cap A = \varnothing,\ (U_2 - \{a\}) \cap B = \varnothing,$$

那么，$U = U_1 \cap U_2$ 是点 a 的邻域；且

$$(U - \{a\}) \cap (A \cup B) = \varnothing.$$

从而 $a \,\overline{\in}\, (A \cup B)'$. 由此便得所要证明的结论.

再证(4). $\forall a \in (A')'$，点 a 的任一邻域 U 包含点 $b \in A'$，$b \neq a$. U 也是 b 的邻域，因而 U 包含 $(A - \{b\})$ 的点. 因此，若 $a \,\overline{\in}\, A$，则 $a \in A'$. 由此即证得(4).

六、闭包与闭集

设 E 是拓扑空间$(X, \mathscr{T})$的一个点集. 集合 $\overline{E} = E \cup E'$ 叫做点集 E 的闭包.

拓扑空间中点集的闭包的定义还是逐字照搬度量空间、欧氏空间中点集的闭包的定义. 并且，完全相一致地易证得下面的性质.

定理 5 设 A，B 是拓扑空间$(X, \mathscr{T})$的两个点集，则下列命题为真：

(1) $\overline{\varnothing} = \varnothing$;

(2) $A \subseteq \overline{A}$;

(3) $\overline{\overline{A}} = \overline{A}$;

(4) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

定义 6 设 E 是拓扑空间$(X, \mathscr{T})$的一个点集.

E 叫做 X 的闭集，如果 $E' \subseteq E$；

E 叫做 X 的自密集，如果 $E \subseteq E'$；

E 叫做 X 的完备集，如果 $E = E'$.

在欧氏空间中闭集、自密集、完备集的例子我们已熟知许多，读者可以在前述诸多拓扑空间中发现更多的例子.

与度量空间的情况一样，容易证得下面的结论.

定理 6 设 E 是拓扑空间$(X, \mathscr{T})$的一个点集. 记 $E^c = X - E$，那么

$$E \text{ 是闭集} \Leftrightarrow E^c \text{ 是开集}.$$

由定理 6 与开集公理便得到

定理 7　设拓扑空间$(X, \mathscr{T})$，则下列命题为真：

(1) X 与空集$\varnothing$是闭集；

(2) 有限个闭集的并集是闭集；

(3) 任意多个闭集的交集是闭集.

由上述论述可以看出：拓扑空间$(X, \mathscr{T})$的任一点集 E 的内部 $\mathrm{Int}E$ 是开集；E 的闭包 $\overline{E}$ 是闭集，从而，

$$E \text{ 是闭集} \Leftrightarrow E = \overline{E}.$$

七、边界

拓扑空间$(X, \mathscr{T})$的点集 E 的余集 $E^c = X - E$ 的内部 $\mathrm{Int}E^c$ 叫做点集 E 的外部，记为 $\mathrm{Ext}E$. $\mathrm{Ext}E$ 的点叫做点集 E 的外点. 容易知道

$$\mathrm{Int}E = (\overline{E^c})^c,\ \mathrm{Ext}E = (\overline{E})^c.$$

集合 $\partial E = \overline{E} - \mathrm{Int}E$ 叫做点集的边界. ∂E 的点叫做点集 E 的边界点. 易知，点集 E 的边界点的任意邻域既与 E 相交，又与 E^c 相交.

八、稠密

定义 7　拓扑空间$(X, \mathscr{T})$的点集 E 说是在空间 X 中处处稠密或稠密，如果 $\overline{E} = X$.

定理 8　拓扑空间$(X, \mathscr{T})$的点集 E 在 X 中处处稠密的充分必要条件是 $\forall U \in \mathscr{T}, U \neq \varnothing$，有 $U \cap E \neq \varnothing$.

证　必要性：假设 $\exists U \in \mathscr{T}, U \neq \varnothing$，使得 $U \cap E = \varnothing$. 则 $\exists a \in U \subseteq X$，由 U 是开集，U 是 a 的一个邻域及 $U \cap E = \varnothing$，知 $a \in E' \cup E = \overline{E}$.

定义 8　拓扑空间$(X, \mathscr{T})$的点集 E 说是在 X 中无处稠密，如果 $\mathrm{Ext}E$ 在 X 中处处稠密.

定理 9　拓扑空间$(X, \mathscr{T})$的点集 E 在 X 中无处稠密的充分必要条件是

$$\forall U \in \mathscr{T}, U \neq \varnothing,\ \exists V \in \mathscr{T}, V \subseteq U, V \neq \varnothing,$$

使得 $V \cap E = \varnothing$.

证　必要性：$\forall U \in \mathscr{T}, U \neq \varnothing$. 由 $\mathrm{Ext}E$ 在 X 中处处稠密，有 $V = U \cap \mathrm{Ext}E \neq \varnothing$. 而 $\mathrm{Ext}E \in \mathscr{T}$，故 $V \in \mathscr{T}, V \subseteq U$，且

$$V \cap E = U \cap \mathrm{Ext}E \cap E = \varnothing.$$

充分性：$\forall a \in X$，a 的任意邻域 U 由题设有非空开集 $V \subseteq U$，使得 $V \cap E = \varnothing$，因而 $\forall b \in V$，有 $b \overline{\in} \overline{E}$，即 $b \in (\overline{E})^c = \mathrm{Ext}E$，故 $a \in (\overline{\mathrm{Ext}\ E})$，即$\mathrm{Ext}E$在

X 中处处稠密.

九、子空间与积空间

拓扑空间$(X, \mathscr{T})$的子空间$(A, \mathscr{T}^*)$是指 A 是 X 的非空子集,且

$$\mathscr{T}^* = \{U \cap A \mid U \in \mathscr{T}\}.$$

拓扑空间$(X_1, \mathscr{T}_1)$与$(X_2, \mathscr{T}_2)$的积空间$(X, \mathscr{T})$是指

$$X = X_1 \times X_2$$

且 $\mathscr{T}$ 的拓扑基 $\mathscr{B} = \{U_\alpha \times V_\beta\}$,其中$\{U_\alpha\}$, $\{V_\beta\}$ 分别是 $\mathscr{T}_1$, $\mathscr{T}_2$ 的拓扑基.

如果 A 是拓扑空间$(X, \mathscr{T})$的一个非空子集. 则易验证集合

$$\mathscr{T}^* = \{U \cap A \mid U \in \mathscr{T}\}$$

是 A 上的一个拓扑,因而$(A, \mathscr{T}^*)$就是拓扑空间$(X, \mathscr{T})$的一个子空间. 如果 A 的子集 B 是 A 的开集,B 不一定是 X 的开集. 但有 X 的开集 V 使得 $V \cap A = B$, 我们有时说B 相对开于 A. 同样可以推知,如果 A 的子集 B 是 A 的闭集,B 不一定是 X 的闭集,但有 X 的闭集 V 使得 $V \cap A = B$, 我们有时说B 相对闭于A.

十、连续映射与拓扑映射

有了邻域的概念,完全可以将度量空间上的连续映射、拓扑映射推广到拓扑空间. 具体地说,只要将本篇第一章第三节的定义 1、定义 2 以及定理 1 中的"度量空间"改写为"拓扑空间";将球形邻域改述为拓扑空间的邻域,就得到拓扑空间的**连续映射**、**连续函数**、**拓扑映射**、**拓扑性质**以及**同胚**的定义,得到定理 1 中的命题(1)、(2)、(3)、(4)是等价命题的结论.

第二节　连　通　性

在几何空间中,图形的连通与否是凭直观感受到的. 这对于任何人来说,都是普通生活中随时可能体验着的概念. 几何学的许多重要结果其实都利用了这一概念. 然而,人们却很少考虑连通性应当怎样严格定义,它的实质是什么. 在数学分析中,曾经描述了区域的连通性. 连通性是重要的拓扑性质. 亦即在拓扑映射下不变的性质. 本节,我们将对拓扑空间的连通性加以严格定义,从而揭示连通性的实质.

一、连通空间

定义 1　拓扑空间$(X, \mathscr{T})$说是**非连通的**,如果

$$\exists F, G \in \mathscr{T}, F \neq \varnothing, G \neq \varnothing, (F \cap \overline{G}) \cup (\overline{F} \cap G) = \varnothing,$$

使得 $X = F \cup G$. 不是非连通的拓扑空间叫做连通空间.

如果拓扑空间 X 的非空点集 A 作为 X 的子空间是连通的，则说 A 是拓扑空间 X 的一个连通集.

我们规定：空集是任何拓扑空间的连通集.

定理 1 设拓扑空间 $(X, \mathscr{T})$，下列命题等价：

(1) X 是连通空间；

(2) 不存在 X 的非空闭集 F 与 G，使得 $X = F \cup G$ 且 $F \cap G \neq \varnothing$；

(3) X 没有既是开集又是闭集的非空真子集.

证 (1)⇒(2)：假若有 X 的非空闭集 F 与 G，使得 $X = F \cup G$ 且 $F \cap G \neq \varnothing$，则 $X = (X - G) \cup (X - F)$，而 $X - G$，$X - F$ 是 X 的非空开集，且 $(X - G) \cap (X - F) = F \cap G = \varnothing$. 这与 X 是连通空间矛盾.

(2)⇒(3)：假设 X 的非空真子集 F 既是开集，又是闭集，那么 $G = X - F$ 是 X 的非空真子集，既是闭集又是开集. 显然有 $F \cap G = \varnothing$，$X = F \cup G$. 这与(2)矛盾.

(3)⇒(1)：假设 X 不是连通空间. 则有 X 的非空开集 F 与 G，使得 $X = F \cup G$ 且 $F \cap G = \varnothing$. 这样，F 与 G 都是 X 的非空真子集. 并且，由于 $F = X - G$，$G = X - F$，F 与 G 都既是开集又是闭集. 这与(3)矛盾.

定义 2 设 A 与 B 是拓扑空间 $(X, \mathscr{T})$ 的两个非空点集，A 与 B 说是互相隔离的，如果 $A \cap B = \varnothing$，并且 A 与 B 都是 X 的子空间 $A \cup B$ 的开集.

定理 2 设 A 与 B 是拓扑空间 $(X, \mathscr{T})$ 的两个非空点集. 则 A 与 B 是互相隔离的充分必要条件是 $\overline{A} \cap B = \overline{B} \cap A = \varnothing$.

证 非空点集 A 与非空点集 B 是互相隔离的

$\Leftrightarrow A \cap B = \varnothing$ 且 A 与 B 都是 $A \cup B$ 的闭集

$\Leftrightarrow A = \overline{A} \cap (A \cup B)$ 且 $B = \overline{B} \cap (A \cup B)$ 且 $A \cap B = \varnothing$

$\Leftrightarrow \overline{A} \cap B = \varnothing$ 且 $\overline{B} \cap A = \varnothing$.

定理 3 设 Y 是拓扑空间 $(X, \mathscr{T})$ 的子集，Y 是连通子集的充分必要条件是 Y 不是 X 的两个互相隔离的子集 A 与 B 的并集.

证 必要性：假设有 $Y = A \cup B$，且 $\overline{A} \cap B = \overline{B} \cap A = \varnothing$. 那么，由 $A \cap B \subseteq \overline{A} \cap B = \varnothing$ 知 $A \cap B = \varnothing$. $\forall a \in A$，由 $\overline{B} \cap A = \varnothing$ 知 $a \bar{\in} \overline{B}$，因而 $\exists a$ 的一个邻域 U_a，使得 $U_a \cap B = \varnothing$，于是 $U_a \subseteq A$. 这样，A 是 X 的开集. 同理 B 也是 X 的开集. 这与 Y 是连通子集矛盾.

充分性：假设 Y 不是连通子集. 则有 $Y = A \cup B$，$A \cap B = \varnothing$，A，B 都是 Y 的既是开集又是闭集的非空真子集. 由于 A 是 Y 的开集. $\forall a \in A$，a 有一个(在 Y 的)邻域含于 A，这个邻域是 Y 的一个开集，因而 $\exists U \in \mathscr{T}$，使得 $U \cap Y$ 是

a 在 Y 的含于 A 的邻域. 即 $a \in U \cap Y \subseteq A$,又 $A \cap B = \varnothing$,因而$(U \cap Y) \cap B = \varnothing$,这便说明 $a \bar{\in} \bar{B}$,因而 $A \cap \bar{B} = \varnothing$. 同理可证 $\bar{A} \cap B = \varnothing$. 于是 A, B 是互相隔离的,矛盾.

定理 4　设 Y 是拓扑空间$(X, \mathscr{T})$的一个连通子集,$\{Y_\alpha\}$是 X 的一族连通子集, $\forall \alpha$, $Y \cap Y_\alpha \neq \varnothing$,那么,$Y \cup (\bigcup_\alpha Y_\alpha)$ 是 X 的连通子集.

证　假设 $(Y \cup (\bigcup_\alpha Y_\alpha)$ 不是 X 的连通子集. 则有 $Y \cup (\bigcup_\alpha Y_\alpha)$ 的既是开集又是闭集的非空真子集 A 与 B,使得

$$Y \cup (\bigcup_\alpha Y_\alpha) = A \cup B \text{ 且 } A \cap B = \varnothing.$$

于是

$$Y = (Y \cap A) \cup (Y \cap B) \text{ 且} (Y \cap A) \cap (Y \cap B) = \varnothing.$$

Y 作为 $Y \cup (\bigcup_\alpha Y_\alpha)$ 的子空间,$Y \cap A$, $Y \cap B$ 是 Y 的两个既是开集又是闭集的真子集. 由 Y 连通,便知它们一个是 Y,一个是$\varnothing$. 不妨,设 $Y \cap A = Y$, $Y \subseteq A$, $Y \cap B = \varnothing$.

同理 $\forall \alpha$, $Y_\alpha \cap A$, $Y_\alpha \cap B$ 一个是 Y_α,一个是$\varnothing$. 但 $Y_\alpha \cap Y \neq \varnothing$, 因而 $Y_\alpha \cap A \neq \varnothing$,故 $Y_\alpha \cap A = Y_\alpha$,而 $Y_\alpha \cap B = \varnothing$. 这样 $Y \cup (\bigcup_\alpha Y_\alpha) \cap B = \varnothing$, 因而 $Y \cup (\bigcup_\alpha Y_\alpha) = A$, $B = \varnothing$, 矛盾.

由定理 4 可知,如果拓扑空间 X 的任意两点都属于某个连通子集,则 X 是连通的.

定理 5　设 Y 是拓扑空间$(X, \mathscr{T})$的连通子集, $Y \subseteq A \subseteq \bar{Y}$, 则 A 是 X 的连通子集.

证　假设 A 有既是开集又是闭集的子集 F 与 G,使得 $A = F \cup G$ 且 $F \cap G = \varnothing$. 我们证明 F、G 中有一者必为空集则可. 由 $Y \subseteq A$, 有

$$Y = Y \cap A = Y \cap (F \cup G) = (Y \cap F) \cup (Y \cap G),$$

$Y \cap F$, $Y \cap G$ 是 Y 的两个既是开集又是闭集的子集,且

$$(Y \cap F) \cap (Y \cap G) = \varnothing,$$

由 Y 是连通子集, $Y \cap F$, $Y \cap G$ 有一者是空集,不妨设 $Y \cap G = \varnothing$. 则 $Y \subseteq F$. 因此,$\bar{Y} \subseteq \bar{F}$. 因为 $A \subseteq \bar{Y}$, 及 F 是 A 的闭集,所以 $A = A \cap \bar{F} = A \cap \bar{Y} = F$. 于是 $G = \varnothing$.

定理 6　设$(X, \mathscr{T})$是连通空间,$(Y, \mathscr{T}_1)$是拓扑空间,连续映射 $f: X \to Y$. 那么 $f(X)$是 Y 的连通集.

证　$f(X)$是 Y 的子空间,也是拓扑空间. 假定 $f(X)$非连通,则有 $f(X)$的

既是开集又是闭集的非空真子集 A, B 使得

$$f(X) = A \cup B \text{ 且 } A \cap B = \varnothing,$$

由 f 是连续映射知，$f^{-1}(A)$, $f^{-1}(B)$是 X 的既是开集又是闭集的非空真子集，使得

$$X = f^{-1}(A) \cup f^{-1}(B) \text{ 且 } f^{-1}(A) \cap f^{-1}(B) = \varnothing,$$

这与$(X, \mathscr{T})$是连通空间矛盾.

定理 6 说明了：**空间的连通性是一个拓扑性质.**

例 1 平庸空间$\{X, \varnothing\}$没有既是开集又是闭集的非空真子集，因而是连通空间.

例 2 至少含 2 个点的离散空间$(X, 2^X)$的独点集既是开集，又是闭集，且是空间的非空真子集，因而离散空间是非连通的.

例 3 W. Sierpinski 空间没有既是开集又是闭集的非空真子集，因而是连通空间.

例 4 拓扑空间的独点集显然是连通集.

例 5 欧氏空间 $\mathbf{E}^1$ 是连通空间.

证 假设 $\mathbf{E}^1$ 非连通，则有 $\mathbf{E}^1$ 的既是开集又是闭集的非空真子集，使得

$$\mathbf{E}^1 = A \cup B \text{ 且 } A \cap B = \varnothing,$$

取 $a \in A$, $b \in B$，不妨令 $a < b$. 记闭区间$[a, b] = I$. 则非空集合 $A \cap I$ 与 $B \cap I$ 是 $\mathbf{E}^1$ 的有界闭集，从而是列紧的. 记数集 $B \cap I$ 的下确界为 λ，由于 $A \cap I$ 与 $B \cap I$ 的距离 $d > 0$，故 $a \neq \lambda$，因此 $a < \lambda$. 故可取满足 $a < x < \lambda$ 且 $\lambda - x < \frac{d}{2}$ 的一个数 x. 这是可以办到的，例如，取正整数 $N_0 > \frac{d}{\lambda - a}$，则 $x = \lambda - \frac{1}{2N_0}d$ 便是满足要求的一个数. 那么，一方面，如果 $x \in A$，则由 $\lambda - x < \frac{d}{2}$ 可推知 $B \cap I$ 与 $A \cap I$ 的距离小于 d，矛盾. 另一方面，如果 $x \in B$，则由 $x < \lambda$，与 λ 是 $B \cap I$ 的下确界矛盾. 故 $x \overline{\in} A \cup B = \mathbf{E}^1$，矛盾. 因此 $\mathbf{E}^1$ 是连通空间.

例 5 实质上严格地论证了我们熟悉的直观：**直线是连通的.**

由于 $\mathbf{E}^1$ 是连通的，$\mathbf{E}^1$ 上的连续函数

$$f(x) = \tan \frac{\pi}{b-a}\left(x - \frac{a+b}{2}\right)$$

的象 $f(\mathbf{E}^1)$也是连通的，亦即区间(a, b)是连通的，再由定理 5 便知半开区间$[a, b)$, $(a, b]$，闭区间$[a, b]$都是连通的. 由 $\mathbf{E}^1$ 与区间同胚也可得此结论.

同样的道理，$\mathbf{E}^1$ 中的区间 I 上的连续函数 f 的象集$f(I)$是 $\mathbf{E}^2$ 的连通集.

定理 7 **设拓扑空间$(X, \mathscr{T})$是连通空间，连续函数 $f: X \to \mathbf{E}^1$. $x_0, x_1 \in X$, $f(x_0) < f(x_1)$. 那么，对于任意满足 $f(x_0) < c < f(x_1)$ 的数 c，存在 $x \in X$，使得 $f(x) = c$.**

证 由于 $\mathbf{E}^1 - \{c\}$ 非连通. 故有 $\mathbf{E}^1 - \{c\}$ 的既是开集又是闭集的非空真子集 A, B，使得

$$\mathbf{E}^1 - \{c\} = A \cup B \text{ 且 } A \cap B = \varnothing,$$

并且 $f(x_0) \in A$, $f(x_1) \in B$.

假设 $c \overline{\in} f(X)$，则

$$\begin{aligned} f(X) &= f(X) \cap (A \cup B \cup \{c\}) \\ &= (f(X) \cap A) \cup (f(X) \cap B), \end{aligned}$$

$f(X) \cap A$, $f(X) \cap B$ 是 $f(X)$的既是开集又是闭集的非空真子集. 这与 $f(X)$ 应是连通集矛盾.

定理 7 是数学分析的中值定理在连通空间的推广.

二、道路连通

定义 3 **$(X, \mathscr{T})$是拓扑空间，a, b 是 X 的点，I 是闭区间$[0, 1]$. 连续映射 $f: I \to X$ 叫做 X 中的连接 a 与 b 的一条道路，如果 $f(0) = a$, $f(1) = b$.**

如果对于 X 的任意两点，都有 X 中的连接它们的道路，就说拓扑空间$(X, \mathscr{T})$是道路连通的.

如果拓扑空间 X 的子空间 A 是道路连通的，则说 A 是 X 的道路连通集.

容易知道，道路连通性是一个拓扑性质. n 维欧氏空间中的区间是道路连通集.

定理 8 **道路连通的拓扑空间是连通空间.**

证 假设拓扑空间$(X, \mathscr{T})$道路连通，但不是连通空间. 则有 X 的既是开集又是闭集的非空真子集 A, B，使得

$$X = A \cup B \text{ 且 } A \cap B = \varnothing,$$

取 $a \in A$, $b \in B$. 因 X 道路连通. 故有连接点 a 与点 b 的一条道路 $f: I \to X$. 这里 I 是闭区间$[0, 1]$. 那么

$$\begin{aligned} f(I) &= f(I) \cap X = f(I) \cap (A \cup B) \\ &= (f(I) \cap A) \cup (f(I) \cap B), \end{aligned}$$

$f(I) \cap A$, $f(I) \cap B$ 便是 $f(I)$的既是开集又是闭集的非空真子集. 这与 $f(I)$ 连通矛盾.

连通的拓扑空间不一定是道路连通的.

例 6 设欧氏空间 $\mathbf{E}^2$ 的点集

$$A = \left\{(x,\ y)\ \middle|\ y = \sin\frac{1}{x},\ 0 \leqslant x \leqslant 1\right\},$$

$$B = \{(0,\ y)\ |-1 \leqslant y \leqslant 1\},$$

由区间$(0,\ 1]$是连通的及函数 $y = \sin\dfrac{1}{x}$ 在$(0,\ 1]$上连续便知 A 是连通集. 因而 $\bar{A} = A \cup B$ 也是连通集.

但 $\bar{A}$ 不是道路连通集. 可以证明,当 $a \in A$, $b \in B$ 时,$\bar{A}$ 中没有连接 a 与 b 的道路.

三、连通分支

定义 4 **拓扑空间$(X,\ \mathscr{T})$的连通集 A 叫做 X 的一个连通分支,如果 X 的任意真包含 A 的子集 B 都是非连通.**

定理 9 **拓扑空间$(X,\ \mathscr{T})$的连通子集是 X 的一个连通分支的子集. X 的连通分支是 X 的闭集.**

证 设 A 是 X 的连通子集,$A \neq \varnothing$, B 是包含 A 的所有连通子集的并集. 则 B 是连通集. 设 X 的子集 C 真包含 B,我们断言,C 非连通. 事实上,假若 C 连通,则 C 也包含 A 的连通子集,因而 $C \subseteq B$, 这与 C 真包含 B 矛盾. 这便证明了 B 是 X 的一个连通分支且 $A \subseteq B$.

又,B 连通,$\bar{B}$ 也连通,B 是连通分支,故 $\bar{B}$ 不能真包含 B. 因而 $B = \bar{B}$. 故 B 是闭集.

类似定义 4,我们可以定义拓扑空间的道路连通分支. 读者不难仿定义 4 自行给出.

第三节　分　离　性

本节介绍满足某些分离性公理的拓扑空间. 由于在开集公理的基础上增加了一些分离性公理,使这样的拓扑空间具有了一些特殊性质,逐渐接近了我们熟知的空间. 由此,也可以从中看到各个分离性公理对拓扑性质的本质性影响.

一、T_0 空间

定义 1 **拓扑空间$(X,\ \mathscr{T})$叫做 T_0 空间,如果$(X,\ \mathscr{T})$满足第零分离公理:拓扑空间中任意两相异点,至少有一个点的一个邻域不包含另一点.**

第零分离公理又叫做Колмогоров(柯尔莫哥洛夫)分离公理，或公理 T_0.

平庸拓扑空间$(X, \mathscr{T})$的任意两点 a, b 都只有一个邻域 X，因而平庸拓扑空间不是 T_0 空间. 至少含有两个点的离散拓扑空间的任意一点 a 总有邻域$\{a\}$不包含另外的点，因而至少含有两个点的离散拓扑空间是 T_0 空间. 度量空间(S, ρ)中的任意两相异点a, b, a 的邻域$O(a, \rho(a, b))$不包含 b，因而，度量空间是 T_0 空间. W. Sierpinski 空间$(X,\mathscr{T})$，这里

$$X = \{0, 1\}, \mathscr{T} = \{X, \varnothing, \{0\}\}$$

虽然点 1 的邻域包含点 0，但是 0 有邻域$\{0\}$不包含点 1，因而，W. Sierpinski 空间$(X, \mathscr{T})$是 T_0 空间.

定理 1　拓扑空间$(X, \mathscr{T})$是 T_0 空间的充分必要条件是 X 的任意两相异独点子集的闭包不相等.

证　必要性：假设 $\exists a, b \in X$，使得$\{\bar{a}\} = \{\bar{b}\}$. 那么，$a \in \{\bar{b}\}$，但 $a \bar{\in} \{b\}$，故 $a \in \{b\}'$，即 a 是$\{b\}$的聚点. 这就是说 a 的任意邻域都包含$\{b\}$的点，亦即包含 b. 同理，b 的任意邻域都包含 a. 这与 X 是 T_0 空间矛盾.

充分性：设 $a, b \in X$, $a \neq b$. 如果点 a 的任意邻域都包含点b，则 b 是$\{a\}$的聚点，因此 $b \in \{\bar{a}\}$，于是$\{b\} \subseteq \{\bar{a}\}$，因而$\{\bar{b}\} \subseteq \{\bar{a}\}$. 同理，如果点 b 的任意邻域都包含点 a，便有 $\{\bar{a}\} \subseteq \{\bar{b}\}$. 因此，假若 X 不是 T_0 空间，便有 $\{\bar{a}\} = \{\bar{b}\}$. 矛盾.

二、T_1 空间

定义 2　拓扑空间$(X, \mathscr{T})$叫做T_1 空间，如果$(X, \mathscr{T})$满足第一分离分理：拓扑空间中任意两相异点，它们中的每一个点都有一个邻域不包含另一个点.

第一分离公理又叫做Fréchet(弗雷歇)分离公理，或公理 T_1.

由定义可知，T_1 空间必是 T_0 空间. 但 T_0 空间不一定是 T_1 空间. 例如，前述的 W. Sierpinski 空间就是不是 T_1 空间的 T_0 空间. 至少含有两个点的离散拓扑空间是 T_1 空间. 度量空间是 T_1 空间.

定理 2　设拓扑空间$(X, \mathscr{T})$，则下列命题等价：

(1) $(X,\mathscr{T})$是 T_1 空间；

(2) X 的独点子集是闭集；

(3) X 的有限子集是闭集；

(4) $\forall a \in X$，包含点 a 的所有的开集的交集是$\{a\}$.

证　(1)⇒(2)：$\forall a \in X$，假设 $b \in X$, $b \neq a$, b 是$\{a\}$ 的聚点. 那么，b 的任意邻域都包含有$\{a\}$的点，亦即包含 a，这与 X 是 T_1 空间矛盾. 因此，$b \bar{\in} \{\bar{a}\}$，从而 $\{\bar{a}\} = \{a\}$, $\{a\}$ 是闭集.

(2)⇒(3)：有限子集是有限个独点子集的并集. 由于独点子集是闭集，因而

有限子集也是闭集.

(3)⇒(4):设$\{G_a\}$是包含点a的所有开集的集合.显然,$a\in\bigcap_a G_a$.假设还有$b\in X$,$b\neq a$,使得$b\in\bigcap_a G_a$,则点a的任意邻域包含b,从而a是b的聚点,$a\in\{\overline{b}\}$,因此,$\{\overline{b}\}\neq\{b\}$,这便与$\{b\}$是闭集矛盾.

(4)⇒(1):假设$(X,\mathscr{T})$不是T_1空间,则X有相异点a,b,使得a的任意邻域都含有b.那么,包含点a的所有的开集的交集必含有点b,这与(4)矛盾.

定理 3　设E是T_1空间$(X,\mathscr{T})$的一个点集,$x\in X$.那么,x是E的聚点的充分必要条件是x的任意邻域都包含E的无穷多个点.

证　充分性是显然的.下证必要性.

任取x的一个邻域U,由x是A的聚点,$\exists x_1\in U\cap(E-\{x\})$;

因X是T_1空间,故有x的邻域V_1使得$x_1\overline{\in}V_1$,记$U_2=U\cap V_1$,则U_2是x的邻域且$x_1\overline{\in}U_2$.由x是E的聚点,$\exists x_2\in U_2\cap(E-\{x\})$;

……

如此下去,便在U中找到E的无穷多个点x_1,x_2,…,x_n,….

定理 4　设E是T_1空间$(X,\mathscr{T})$的一个点集,则$(E')'\subseteq E'$.

证　$\forall x\in(E')'$,x是E'的聚点.任取点x的邻域U,$\exists a\in U\cap(E'-\{x\})$.因$X$是$T_1$空间,故有$a$的邻域$V$,使得$x\overline{\in}V$.记$W=U\cap V$,则$W$是$a$的一个邻域,且$x\overline{\in}W\subseteq U$.注意到$a$是$E$的聚点,故$\exists b\in W\cap(E-\{a\})$.因此,$b\in U\cap(E-\{x\})$,即$x$是$E$的聚点,$x\in E'$.

三、T_2 空间

定义 3　拓扑空间$(X,\mathscr{T})$叫做T_2空间,如果$(X,\mathscr{T})$满足第二分离公理:拓扑空间中任意两相异点有不相交的邻域.

第二分离公理通常叫做Hausdorff(豪斯多夫)分离公理,或公理T_2.T_2空间常被叫做Hausdorff空间.

由定义可知,T_2空间必是T_1空间.但T_1空间不一定是T_2空间.下举一例.

例 1　设X是全体正整数的集合.

$$\mathscr{T}=\{A\mid A=X-B,\ B\text{是}X\text{的有限子集}\}\cup\{\varnothing\},$$

则$(X,\mathscr{T})$是一个拓扑空间.

$\forall n\in X$,由于$X-\{n\}$是开集,故$\{n\}$是闭集.由此便知X是T_1空间.

设A_m,A_n分别是X的点m,n的邻域.据$\mathscr{T}$的定义,有X的有限子集B_m.B_n使得$A_m=X-B_m$,$A_n=X-B_n$.于是

$$A_m \cap A_n = (X - B_m) \cap (X - B_n) = X - (B_m \cup B_n),$$

因此，$A_m \cap A_n$ 是无限集，特别地，$A_m \cap A_n \neq \varnothing$. 由此便知 X 不是 T_2 空间.

定理 5 **拓扑空间**$(X, \mathscr{T})$**是** T_2 **空间的充分必要条件是** $\forall a, b \in X, a \neq b$, $\exists U \in \mathscr{T}$, **使得** $a \in U$ **且** $b \bar{\in} \bar{U}$.

证 必要性：$\forall a, b \in X, a \neq b$, 由 X 是 T_2 空间，$\exists U, V \in \mathscr{T}$, 使得

$$a \in U, b \in V, U \cap V = \varnothing.$$

由此可知，b 的邻域 V 不包含点集 U 的点，故 $b \bar{\in} \bar{U}$.

充分性：由题设，$\forall a, b \in X, a \neq b$, 点 a 的邻域 U 与点 b 的邻域 $X - \bar{U}$ 不相交. 因此，X 是 T_2 空间.

定理 6 T_2 **空间的子空间与积空间都是** T_2 **空间.**

证明留给读者.

定理 7 **设拓扑空间**$(X, \mathscr{T})$**是** T_2 **空间，则** X **的收敛序列**$\{x_n\}$**有惟一的极限.**

证 假设 a, b 是$\{x_n\}$的两个极限，$a \neq b$. 由 X 是 T_2 空间，a, b 分别有邻域 U, V 使得 $U \cap V = \varnothing$.

由 $x_n \to a$ 知，$\exists$ 正整数 N_1，使得当 $n > N_1$ 时，$x_n \in U$；

由 $x_n \to b$ 知，$\exists$ 正整数 N_2，使得当 $n > N_2$ 时，$x_n \in V$.

那么，当 $n > N' = \max\{N_1, N_2\}$ 时，$x_n \in U \cap V$. 矛盾.

四、正则空间与 T_3 空间

定义 4 **拓扑空间**$(X, \mathscr{T})$**叫做正则空间，如果**$(X, \mathscr{T})$**满足第三分离公理：拓扑空间的任意一点与不包含这个点的任意闭集有不相交的邻域.**

第三分离公理又叫做Vietories(维多里斯)分离公理，或公理 T_3.

应当指出，正则空间不一定是 T_2 空间，甚至不必是 T_1 空间，乃至不必是 T_0 空间.

例 2 设 $X = \{a, b, c\}$, $\mathscr{T} = \{X, \varnothing, \{a\}, \{b, c\}\}$, 则$(X, \mathscr{T})$是一个拓扑空间，它有且只有 4 个开集，而且这 4 个开集恰好是它的全部闭集. 易验证 X 是正则空间. 由于点 b 与点 c 总是有共同的邻域 X 或$\{b, c\}$，因此，X 不是 T_0 空间，自然就不是 T_1 空间，不是 T_2 空间了.

反之，T_2 空间也不一定是正则空间.

例 3 2 维欧氏空间 $(\mathbf{R}^2, d)$. $\forall p = (a, b) \in \mathbf{R}^2$, 定义点集

$$L_p = \{(x, y) \mid x = a, y \in \mathbf{R}, y \neq b\},$$

$\forall r > 0$, 定义点集

$$U(p, r) = O(p, r) - L_p,$$

令
$$\mathscr{B} = \{U(p, r) \mid p \in \mathbf{R}^2, r > 0\},$$

则可验证 $\mathscr{B}$ 是一个拓扑空间 $(\mathbf{R}^2, \mathscr{T})$ 的拓扑基. 在这个拓扑空间中, $U(p, r)$ 是点 p 的邻域. $\forall p, q \in \mathbf{R}^2$, 有

$$U\left(p, \frac{1}{2}d(p, q)\right) \cap U\left(q, \frac{1}{2}d(p, q)\right) = \varnothing,$$

因此 $(\mathbf{R}^2, \mathscr{T})$ 是 T_2 空间. $\forall p \in \mathrm{R}^2$, 再令

$$F = \mathbf{R}^2 - U(p, r),$$

则 F 是 $(\mathbf{R}^2, \mathscr{T})$ 的闭集, 且 $p \overline{\in} F$. 任取 F 的邻域 V, 任取点 p 的邻域, 它必包含一个 $U(p, r_1)$,

$$O(p, r_1) \cap L_p \neq \varnothing,$$

故 $\exists q \in O(p, r_1) \cap L_p$, 因而 $q \in F \subseteq V$. 由 $\mathscr{B}$ 是拓扑基, $\exists U(q, r_2) \subseteq V$.

显然
$$U(q, r_2) \cap U(p, r_1) \neq \varnothing,$$

因而
$$V \cap U(p, r_1) \neq \varnothing,$$

因此, $(\mathbf{R}^2, \mathscr{T})$ 不是正则空间.

定义 5 **拓扑空间 $(X, \mathscr{T})$ 叫做 T_3 空间, 如果 $(X, \mathscr{T})$ 是正则空间, 并且是 T_1 空间.**

定理 8 **T_3 空间必是 T_2 空间.**

证 因为 T_3 空间是 T_1 空间, 故独点集都是闭集. 因此, T_3 空间中的两个相异点 a, b, 由 T_3 空间是正则空间, 存在点 a 的邻域 U 与闭集 $\{b\}$ 的邻域 V, 使得 $U \cap V = \varnothing$. 而 V 也是点 b 的邻域. 因此这个空间也是 T_2 空间.

由例 3 可知, T_2 空间不必是正则空间, 因而也不必是 T_3 空间.

度量空间是 T_3 空间. 这是因为度量空间 (S, ρ) 的闭集 F 与点 $a \overline{\in} F$ 的距离 $d = \rho(a, F) > 0$, 因此, 点 a 的邻域 $O\left(a, \frac{d}{2}\right)$ 与 F 的邻域

$$O\left(F, \frac{d}{2}\right) = \left\{x \,\middle|\, x \in S, \rho(x, F) < \frac{d}{2}\right\}$$

不相交. 因而 (S, ρ) 是正则空间. 又度量空间显然是 T_1 空间. 因此, 度量空间是 T_3 空间.

定理 9 **拓扑空间 $(X, \mathscr{T})$ 是正则空间的充分必要条件是 $\forall a \in X$, $\forall a$ 的邻域 U, $\exists a$ 的邻域 V, 使得 $\overline{V} \subseteq U$.**

证 必要性: $\forall a$ 的邻域 U, 令 $F = X - U$, 则 F 是闭集, 且 $a \overline{\in} F$. 因 X 是

正则空间，故有 a 的邻域 V 与 F 的邻域 W，使得

$$V \cap W = \varnothing.$$

因 W 是开集，故 $W^c = X - W$ 是闭集. 由 $V \cap W = \varnothing$. 有 $V \subseteq W^c$，从而 $\overline{V} \subseteq W^c$. 再因 $W^c \subseteq X - F = U$，所以 $\overline{V} \subseteq U$.

充分性：$\forall a \in X$，任取闭集 F，$a \bar{\in} F$. 那么 $U = X - F$ 是 a 的一个邻域. 由题设，有 a 的邻域 V，使得 $\overline{V} \subseteq U$. 令 $W = X - \overline{V}$，则 W 是开集. 因 $\overline{V} \subseteq U$，故 $X - \overline{V} \subseteq X - U$，那么 $W \supseteq F$. 因而 W 是闭集 F 的邻域. 由 $W \cap V = \varnothing$ 便知 X 满足第三分离公理，所以它是正则空间.

定理 10　正则空间的子空间是正则空间.

证　设 A 是正则空间 X 的一个子空间. $\forall x \in A$，设 G 是 A 的闭集且 $x \bar{\in} G$. 那么，有 X 的闭集 F，使得 $G = A \cap F$. 由 X 是正则空间，有 x 在 X 的邻域 V_1 与 F 在 X 的邻域 V_2，使得 $V_1 \cap V_2 = \varnothing$. 那么，$V_1 \cap A$ 是 x 在 A 的邻域，$V_2 \cap A$ 是 G 在 A 的邻域，且 $(V_1 \cap A) \cap (V_2 \cap A) = \varnothing$. 故 A 是正则空间.

五、正规空间与 T_4 空间

定义 6　拓扑空间 $(X, \mathscr{T})$ 叫做正规空间，如果 $(X, \mathscr{T})$ 满足第四分离分理：拓扑空间任意两个不相交的闭集，有不相交的邻域.

第四分离公理又叫做铁策分离公理，或公理 T_4.

正规空间不必是 T_2 空间. 例如，本节例 2 所示的拓扑空间 $X = \{a, b, c\}$ 有 4 个开集：

$$X, \varnothing, \{a\}, \{b, c\}.$$

它们也恰是 X 的 4 个闭集. 由定义可以验证，X 不仅是正则空间，也是正规空间. 但 X 不是 T_2 空间，也不是 T_1 空间，不是 T_0 空间.

T_2 空间不必是正规空间.

正则空间不必是正规空间.

正规空间不必是正则空间.

以上结论都可以通过一些反例说明. 这是不再赘述，读者可以思考一些例证，或者参考篇末所列的文献.

定义 7　拓扑空间 $(X, \mathscr{T})$ 叫做 T_4 空间，如果 $(X, \mathscr{T})$ 是正规空间，并且是 T_1 空间.

定理 11　T_4 空间是 T_3 空间.

证　设 X 是 T_4 空间，$\forall x \in X$，由 T_4 空间是 T_1 空间，$\{x\}$ 是 X 的闭集. 设 F 是 X 的不包含点 x 的闭集，那么，$\{x\} \cap F = \varnothing$. 由 T_4 空间是正规空间，F 与 $\{x\}$ 有不相交的邻域，但 $\{x\}$ 的邻域就是点 x 的邻域. 因此，X 是 T_3 空间.

定理 12　度量空间(S, ρ)是 T_4 空间.

证　设 F, G 是 S 的两个不相交的非空闭集. 作

$$U = \bigcup_{x \in F} O\left(x, \frac{\rho_x}{2}\right), \quad 其中 \rho_x = \rho(x, G);$$

$$V = \bigcup_{y \in G} O\left(y, \frac{\rho_y}{2}\right), \quad 其中 \rho_y = \rho(y, F).$$

因为 F, G 是闭集且 $F \cap G = \varnothing$. 故 $\rho_x > 0$, $\rho_y > 0$. U, V 分别是 F, G 的邻域. 故只须证 $U \cap V = \varnothing$ 则可.

假设 $a \in U \cap V$, 由 U, V 的定义, $\exists x \in F$, $y \in G$, 使得

$$\rho(x, a) < \frac{\rho_x}{2}, \rho(y, a) < \frac{\rho_y}{2}.$$

若 $\rho_x \geqslant \rho_y$, 则

$$\rho_x = \rho(x, G) \leqslant \rho(x, y) \leqslant \rho(x, a) + \rho(y, a) < \frac{\rho_x}{2} + \frac{\rho_y}{2} \leqslant \rho_x, 矛盾.$$

若 $\rho_x \geqslant \rho_y$, 则

$$\rho_y = \rho(x, F) \leqslant \rho(y, x) \leqslant \rho(y, a) + \rho(x, a) < \frac{\rho_y}{2} + \frac{\rho_x}{2} \leqslant \rho_y, 矛盾.$$

故 $U \cap V = \varnothing$. S 是正规空间. 又, 度量空间是 T_1 空间, 因而 S 是 T_4 空间.

应当指出, 定理 11 的逆命题不真. 也就是说, T_3 空间不必是 T_4 空间. 这样的例证这里就不多谈了. 读者可参考书末的文献.

定理 13　拓扑空间$(X, \mathscr{T})$是正规空间的充分必要条件是 $\forall$ 闭集 $F \subseteq X$, $\forall F$ 的邻域 U, $\exists F$ 的邻域 V, 使得 $\overline{V} \subseteq U$.

证　必要性: 设 F 的邻域 U, 令 $U^c = X - U$, 则 U^c 是闭集, 且 $F \cap U^c = \varnothing$. 因 X 是正规空间, 故有 F 的邻域 V 与 U^c 的邻域 W, 使得

$$V \cap W = \varnothing.$$

因 W 是开集, 故 $W^c = X - W$ 是闭集. 由 $V \cap W = \varnothing$, 有 $V \subseteq W^c$, 从而 $\overline{V} \subseteq W^c$. 再因 $W^c \subseteq X - U^c = U$, 所以 $\overline{V} \subseteq U$.

充分性: 设 F, G 是 X 的两个不相交的闭集. $G^c = X - G$ 是 F 的一个邻域. 由题设, 有 F 的邻域 V, 使得 $\overline{V} \subseteq G^c$, 那么 $\overline{V}^c$ 是 G 的邻域. 且 $V \cap \overline{V}^c = \varnothing$, 故 X 是正规空间.

定理 14　设拓扑空间$(X, \mathscr{T})$是正规空间, A, B 是 X 的两个不相交的闭集. 那么, 存在连续函数 $f: X \to [0, 1]$, 使得

$$f(x)=\begin{cases}0, & \text{当 } x\in A \text{ 时;}\\ 1, & \text{当 } x\in B \text{ 时.}\end{cases}$$

定理 14 叫做Урысон(乌里松)引理. 其证明较复杂,这里略去. 读者可参阅篇末文献.

一般的,如果 A, B 是拓扑空间 X 的不相交的两个非空点集,连续函数 $f:X\to[0,1]$,使得 $f(A)=\{0\}$, $f(B)=\{1\}$,则 f 叫作 X 上关于点集偶 (A,B) 的 **Урысон 特征函数**.

Урысон 引理告诉我们:正规空间的任意两个不相交的闭集 A 与 B,存在关于闭集偶 (A,B) 的 Урысон 特征函数.

Урысон 引理的逆命题也成立.

定理 15　拓扑空间 $(X,\mathscr{T})$ 是正规空间的充分必要条件是 X 的任意两个不相交的闭集 A, B,存在关于闭集偶 (A,B) 的 Урысон 特征函数.

证　必要性:由 Урысон 引理而知. 下证充分性:

设 f 是关于闭集偶 (A,B) 的一个 Урысон 特征函数. 取 $[0,1]$ 的两个开集 $F=\left[0,\dfrac{1}{4}\right)$, $G=\left(\dfrac{3}{4},1\right]$. 令 $U=f^{-1}(F)$, $V=f^{-1}(G)$,显然 $A\subseteq U$, $B\subseteq V$,且 $U\cap V=\varnothing$. 故 X 是正规空间.

从本节的论述可以看到,考察满足某些公理的拓扑空间,对拓扑空间适当增加一些限制,能使我们逐步深刻地揭示拓扑空间的性质,能使我们从欧氏空间、度量空间逐步过渡、深入地认识和了解拓扑空间.

第四节　可　数　性

一、可数性公理

本节介绍分别满足两个可数性公理的拓扑空间. 由于拓扑空间的开集、闭集、闭包、收敛序列与邻域都是密切相关的概念,因此,满足可数性公理会对这些概念产生重要影响. 我们所说的两个可数性公理是:

第一可数性公理:拓扑空间的每一点处都有可数的邻域基;

第二可数性公理:拓扑空间有可数的拓扑基.

第一可数性公理简称**公理 A_1**,第二可数性公理简称**公理 A_2**.

显然,如果拓扑空间满足公理 A_2,就一定满足公理 A_1.

n 维欧氏空间、列紧度量空间、有限拓扑空间都有可数拓扑基,因而满足公理 A_2. 度量空间的每一点处的以该点为中心、正有理数为半径的球形邻域族是可数邻域基,因而度量空间满足公理 A_1.

下面举出不满足公理 A_1 与不满足公理 A_2 但满足公理 A_1 的拓扑空间的例.

例 1 设 $X = \{x \mid 0 \leqslant x \leqslant 1\}$, $\mathscr{T} = \{B \mid B = A^c = X - A, A \subseteq X, A$ 至多可数$\} \cup \{\varnothing\}$, 则$(X, \mathscr{T})$是拓扑空间.

下面说明$(X, \mathscr{T})$不满足公理 A_1.

假设 X 的每个点处都有可数邻域基. 记点 x 处的邻域基为$\{A_n^c\}$, A_n 是 X 的至多可数子集. 因 X 不可数,故 $\exists a \in X$, $a \neq x$, $a \,\bar{\in}\, \bigcup\limits_{n=1}^{\infty} A_n$. 因此,

$$x \in \{a\}^c = X - \{a\},$$

而 $X - \{a\} \in \mathscr{T}$, 故$\{a\}^c$ 是 x 的邻域. 但$\{a\}^c$ 不包含任何一个 A_n^c,矛盾.

例 2 $\mathbf{R}$ 是实数集,离散拓扑空间$(\mathbf{R}, 2^{\mathbf{R}})$.

这个空间的每一个子集既是开集又是闭集. 它的拓扑基必须包含所有的独点集,因而不可数. 但独点集$\{x\}$便是点 x 处的一个可数的邻域基. 因此,拓扑空间$(\mathbf{R}, 2^{\mathbf{R}})$满足公理 A_1,但不满足公理 A_2.

二、满足第一可数性公理的拓扑空间

定理 1 **满足公理 A_1 的拓扑空间$(X, \mathscr{T})$,其子空间 $(A, A \cap \mathscr{T})$ 也满足公理 A_1.**

证 $\forall x \in A$,由 $x \in X$, 在点 x 处有点 x 在 X 中的可数邻域基$\{U_n\}$,那么,$\{A \cap U_n\}$便是在点 x 处,x 在 A 中的可数邻域基.

定理 2 **满足公理 A_1 的拓扑空间 X 与 Y 的积空间 $X \times Y$ 也满足公理 A_1.**

证 $\forall (x, y) \in X \times Y$, x, y 分别在 X, Y 中有可数邻域基 $\{U_n\}$, $\{V_n\}$. 那么

$$\{U_i \times V_j \mid i = 1, 2, \cdots; j = 1, 2, \cdots\}$$

便是点(x, y)处的可数邻域基.

定理 3 **设拓扑空间$(X, \mathscr{T})$满足公理 A_1,那么, $\forall x \in X$, 存在 x 处的可数邻域基 W_n,使得 $W_{n+1} \subseteq W_n \quad (n = 1, 2, 3, \cdots)$.**

证 设$\{V_n\}$是点 x 处的可数邻域基. 记

$$W_1 = V_1, \; W_n = \bigcap_{k=1}^{n} V_k (n = 2, 3, \cdots),$$

则得结论.

定理 4 **设拓扑空间$(X, \mathscr{T})$是满足公理 A_1 的 T_1 空间, $x \in X$. 那么,存在 x 处的可数邻域基 W_n,使得 $W_{n+1} \subseteq W_n \quad (n = 1, 2, 3, \cdots)$, 并且**

$$\bigcap_{n=1}^{\infty} W_n = \{x\}.$$

证 由定理3,有 x 的可数邻域基 W_n,使得 $W_{n+1} \subseteq W_n \quad (n=1, 2, 3, \cdots)$. 显然,$x \in \bigcap_{n=1}^{\infty} W_n$. 设 $y \in X$, $y \neq x$. 由 X 是 T_1 空间知 $\{y\}$ 是闭集. 从而 x 不是 $\{y\}$ 的聚点. 故有 x 的一个邻域 U 不包含 y. 由 W_n 是点 x 处的邻域基,存在正整数 N,使得 $W_N \subseteq U$. 因此,$y \overline{\in} W_n$.

设 E 是拓扑空间 $(X, \mathscr{T})$ 是点集. 现在我们讨论 E 的聚点与 E 的收敛序列有关的以下 3 个命题:

命题(1) x 是 E 的聚点;

命题(2) 有 $E-\{x\}$ 的序列 $\{x_n\}$ 收敛于 x;

命题(3) 有 $E-\{x\}$ 的序列 $\{x_n\}$ 收敛于 x,其中 $i \neq j$ 时有 $x_i \neq x_j$.

对此 3 个命题的联系,可得以下结论:

结论 1 对于任意拓扑空间,均有命题(3)⇒命题(2)⇒命题(1).

证 这是显然的.

结论 2 对于满足公理 A_1 的拓扑空间,有命题(1)⇒命题(2).

证 设 $\{W_n\}$ 是点 x 处的一个可数邻域基. x 是 E 的聚点,故对任意正整数 n, $\exists x_n \in W_n \cap (E-\{x\})$, 这样,便得所要求的序列 $\{x_n\}$.

结论 3 存在不满足公理 A_1 的 T_1 空间,使得命题(1)不蕴含命题(2).

本节例 1 的拓扑空间 $(X, \mathscr{T})$ 是不满足公理 A_1 的 T_1 空间. 其中, $X=[0, 1]$, $\mathscr{T}=\{B \mid B=A^c=X-A, A\subseteq X, A \text{ 至多可数}\} \cup \{\varnothing\}$. 如果 E 是 X 的至多可数的子集, $\forall x \in X$, 则因

$$(E-\{x\})^c \cap (E-\{x\}) = \varnothing,$$

而 $(E-\{x\})^c$ 是 x 的邻域,知 x 不是 E 的聚点;如果 E 是 X 的不可数的子集, $\forall x \in X$, x 的邻域 A^c 与 E 的交集 $A^c \cap E$ 不可数,故 x 是 E 的聚点,从而 $\overline{E}=X$.

如果 $\{x_n\}$ 是 X 的收敛序列,其极限为 x,则存在正整数 N,使得 $n>N$ 时, $x_n=x$. 否则的话,点 x 的邻域 $(X-\{x_n\}) \cup \{x\}$ 不包含序列 $\{x_n\}$ 的点,这与 $\{x_n\}$ 收敛于 x 矛盾.

现设 E 是 X 的不可数子集. 则 $\forall x \in X$, x 是 E 的聚点. 但 $E-\{x\}$ 中没有收敛到 x 的序列. 这便说明了命题(1)不蕴含命题(2).

结论 4 对于 T_1 空间,有命题(2)⇒命题(3).

证 设 $A-\{x\}$ 的序列 $\{x_n\}$ 收敛于 x,那么, $\forall n$, $x_n \neq x$. 设序列 $\{x_n\}$ 中的点的全体的集合为 E,则 E 不可能是有限集,否则的话,由于在 T_1 空间中,有限点集是闭集, $X-E$ 便是点 x 的一个邻域,但 $X-E$ 却不包含序列 $\{x_n\}$ 的点. 这与 $\{x_n\}$ 收敛于 x 矛盾. 因此,E 是无限集. 因此,序列 $\{x_n\}$ 便有每两项都不相同的

子列收敛于 x.

结论 5　存在满足公理 A_2 而非 T_1 空间的拓扑空间,使得命题(2)不蕴含命题(3).

设集合　$X=\{a, b, c\}$, $\mathscr{T}=\{X, \varnothing, \{a, b\}, \{a, c\}, \{a\}\}$,则拓扑空间 $(X, \mathscr{T})$ 是 T_0 空间,满足公理 A_2,但非 T_1 空间.

令 $x_n=a$,则常项序列$\{x_n\}$收敛,a, b, c 都是$\{x_n\}$的极限.由此可见命题(2)不蕴含命题(3).

结论 6　对于满足公理 A_1 的 T_1 空间,有命题(1)$\Leftrightarrow$命题(2)$\Leftrightarrow$命题(3).

证　由结论 1,结论 2,结论 4 便知.

最后,我们指出,当映射的定义空间满足第一可数性公理时,函数的连续性可以用序列的收敛性来刻画.

定理 5　设 X, Y 是拓扑空间,X 满足公理 A_1,映射 $f:X\to Y$.那么,f 是连续映射的充分必要条件是 $\forall x\in X$,对于 X 中收敛于 x 的序列$\{x_n\}$,都有$\{f(x_n)\}$收敛于 $f(x)$.

证　必要性是显然的,事实上,对于任意拓扑空间(不必满足公理 A_1)都是成立的,下证充分性.

假设 f 在点 $x\in X$,则 $\exists f(x)$ 的一个邻域 U,使得 x 不是 $f^{-1}(U)$ 的内点.由 X 满足公理 A_1,在点 X 处有满足 $W_{n+1}\subseteq W_n$ 的可数邻域基$\{W_n\}$.由 x 不是 $f^{-1}(U)$的内点.$\forall n$,均有 $W_n\not\subseteq f^{-1}(U)$.因此,$\exists x_n\in W_n$,使得 $x_n\overline{\in} f^{-1}(U)$.这样,便有 X 中收敛于 x 的序列$\{x_n\}$,使得点$\{f(x_n)\}$不收敛于 $f(x)$.这与题设矛盾.

三、覆盖

在讨论满足第二可数性公理的拓扑空间之前,我们有必要将度量空间关于覆盖的概念推广到拓扑空间.这是容易直观理解,容易办到的.具体地说,只要将本篇第一章第四节中的定义 4 的度量空间改述为拓扑空间,其余逐句照搬,便得到拓扑空间中关于覆盖、开覆盖、闭覆盖、可数覆盖、有限覆盖、子覆盖的概念.这里不再重复赘述.

四、满足第二可数性公理的拓扑空间

定理 6　满足公理 A_2 的拓扑空间$(X, \mathscr{T})$,其子空间 $(A, A\cap\mathscr{T})$ 也满足公理 A_2.

证　如果$\{W_n\}$是 X 的可数拓扑基,则 $\{A\cap W_n\}$ 是 A 的可数拓扑基.

定理 7　满足公理 A_2 的拓扑空间 X 与 Y 的积空间 $X\times Y$ 也满足公理 A_2.

证　如果$\{U_n\}$, $\{V_n\}$分别是 X 与 Y 的可数拓扑基,则

$$\{U_i \times V_j \mid i = 1, 2, \cdots; j = 1, 2, \cdots\}$$

是 $X \times Y$ 的可数拓扑基.

定理 8 (Lindelof(林德略夫)定理),**满足第二可数性公理的拓扑空间的任何开覆盖都有一个可数子覆盖.**

证 设拓扑空间$(X, \mathscr{T})$满足公理 A_2, $\mathscr{B}$ 是 X 的一个可数拓扑基. 设 $\mathscr{M} = \{U_\alpha\}$ 是 X 的一个开覆盖. 记

$$\mathscr{U} = \{V \mid V \in \mathscr{B}, \exists U \in \mathscr{M}, \text{使得 } V \subseteq U\}.$$

则 $\mathscr{U} \subseteq \mathscr{B}$, $\mathscr{U}$ 至多可数. 不妨记

$$\mathscr{U} = \{V_1, V_2, \cdots\}$$

$\forall n$,任取一个满足 $V_n \subseteq U_n$ 的 $U_n \in \mathscr{M}$. 记

$$\mathscr{M}_1 = \{U_1, U_2, \cdots\}.$$

我们断言,$\mathscr{M}_1$ 是 X 的一个可数开覆盖. 事实上 $\forall x \in X$, 由 $\mathscr{M}$ 是 X 的覆盖, $\exists U_\alpha \in \mathscr{M}$,使得 $x \in U_\alpha$. $\mathscr{B}$ 是拓扑基. 故存在 $V \in \mathscr{B}$,使得 $x \in V$,且 $V \subseteq U_\alpha$. 由 $\mathscr{U}$ 的定义, $V \in \mathscr{U}$. 不妨,记 $V = V_k$,则 $x \in U_k$.

五、可分空间

定义 1 有可数稠密子集的拓扑空间叫做可分空间,或说是可分的.

定理 9 满足第二可数性公理的拓扑空间是可分空间.

证 设拓扑空间$(X, \mathscr{T})$满足公理 A_2, U_n 是 X 的一个可数拓扑基. 在每个 U_n 中任取且只取一点 x_n,记 A 是取出的这些点的集合. 显然,集合 A 至多可数. 我们断言,A 在 X 中稠密. 事实上, $\forall x \in X$, 设 V 是 x 的任一邻域. 则有某个 $U_n \subseteq U$ 且 $x \in U_n$,因 $x_n \in U_n \cap A$,故 $V \cap A \neq \varnothing$,因而 $x \in \bar{A}$,故 $\bar{A} = X$.

应当指出,定理 9 的逆命题不真. 可分空间不必是满足第二可数性公理的拓扑空间.

例 3 设集合 X 是实数集, $\mathscr{T} = \{A \mid A \subseteq X, A \text{不包含非零有理数}, 0 \in A\} \cup \{X\} \cup \{\varnothing\}$.

易知$(X, \mathscr{T})$是拓扑空间.

有理数集 $\mathbf{Q}$ 是 X 的可数子集. $\forall x \in X$,若 $x \bar{\in} \mathbf{Q}$, 则 x 的任意邻域都包含 $\mathbf{Q}$ 的点 0,因而 $x \in \mathbf{Q}'$,于是$\overline{\mathbf{Q}} = X$. X 是可分空间. 而 X 显然没有可数拓扑基.

第五节 紧 致 性

紧致性是一个十分重要的拓扑性质,在度量空间中,我们已对它有了初步认

识.本节将对它进行较细致的讨论.

一、紧致空间

定义 1 **设 E 是拓扑空间$(X, \mathscr{T})$的点集,如果 E 的任意开覆盖都有一个有限子覆盖,则 E 叫做是空间 X 的一个紧致集,或简称紧集.如果 X 是紧致集,则说 X 是一个紧致拓扑空间,或简称紧致空间,或紧空间.**

例 1 拓扑空间的任意有限集都是紧致集.

例 2 如果 X 是无限集,离散拓扑空间$(X, 2^X)$不是紧致空间.X 的无限子集不是紧致集.这是因为每个独点集都是开集.

例 3 n 维欧氏空间 $\mathbf{E}^n$ 中的有界闭区间是紧致集.这由数学分析中著名的有限覆盖定理可知.

例 4 直线 $\mathbf{E}^1$ 不是紧致空间.因为若记开区间 $(-n, n)=U_n$,则 $\mathscr{M}=\{U_n\}$ 是 $\mathbf{E}^1$ 的一个开覆盖,显然没有有限子覆盖.

定理 1 **紧致空间的闭集是紧致集.**

证 设 E 是紧致空间$(X, \mathscr{T})$的一个闭集.$\mathscr{M}$ 是 E 的一个开覆盖.因 $E^c=X-E\in\mathscr{T}$,故 $\mathscr{M}\cup\{E^c\}$ 是 X 的一个开覆盖.由 X 是紧致空间.$\mathscr{M}\cup\{E^c\}$ 有一个有限子覆盖 $\mathscr{U}$,那么 $\mathscr{U}-\{E^c\}$ 便是 E 的开覆盖 $\mathscr{M}$ 的一个有限子覆盖.

以后,我们要提到术语"有限交性质".设非空集合 S,$\{F_\alpha\}$是 S 的某些子集的集合.所谓集合族$\{F_\alpha\}$具有有限交性质,是指$\{F_\alpha\}$中任意有限个元素的交集都非空集.

定理 2 **拓扑空间$(X, \mathscr{T})$是紧致空间的充分必要条件 X 的任意具有有限交性质的闭集族$\{F_\alpha\}$有 $\bigcap\limits_\alpha F_\alpha\neq\varnothing$.**

证 必要性:假设 $\bigcap\limits_\alpha F_\alpha=\varnothing$,则

$$\bigcap_\alpha F_\alpha^c=(\bigcap_\alpha F_\alpha)^c=X-\bigcap_\alpha F_\alpha=X.$$

故$\{F_\alpha^c\}$是 X 的一个开覆盖.由 X 是紧致空间,有一个有限覆盖

$$\{F_1^c, F_2^c, \cdots, F_n^c\}.$$

这样的话,

$$\bigcap_{k=1}^n F_k=\bigcap_{k=1}^n(X-F_k^c)=X-\bigcup_{k=1}^n F_k^c=\varnothing,$$

与题设$\{F_\alpha\}$具有有限交矛盾.

充分性:假设 X 不是紧致空间.则有 X 的一个开覆盖 $\mathscr{M}=\{U_\alpha\}$ 没有有限子覆盖.记 $F_\alpha=X-U_\alpha=U_\alpha^c$,$F_\alpha^c=X-F_\alpha=U_\alpha$,则 F_α 是闭集.

我们断言,闭集族$\{F_\alpha\}$具有有限交性质.事实上,假设有

$$\bigcap_{k=1}^{n} F_k = \varnothing,$$

则

$$\bigcup_{k=1}^{n} U_k = X,$$

这与 $\mathscr{M}$ 没有有限子覆盖矛盾.

再由 $\mathscr{M}$ 是 X 的覆盖,有

$$\bigcap_{\alpha} F_\alpha = \bigcap_{\alpha} U_\alpha^c = X - \bigcup_{\alpha} U_\alpha = \varnothing,$$

这便与题设矛盾.

二、列紧空间

定义 2 **设 E 是拓扑空间$(X, \mathscr{T})$的点集,如果 E 的每个无穷子集都有一个(属于 E 的)聚点,则 E 叫做是空间 X 的一个列紧集. 如果 X 是列紧集,则说 X 是一个列紧拓扑空间,或简称列紧空间.**

例 5 拓扑空间的任意有限集都是列紧集.

例 6 列紧空间的闭集是列紧集.

例 7 有限个列紧集的并集是列紧集.

定理 3 **紧致空间是列紧空间.**

证 设 A 是紧致空间$(X, \mathscr{T})$的一个无穷点集. 假设 A 无聚点. 则 $\forall x \in X$, $\exists x$ 的邻域 U_x,使得

$$U_x \cap (A - \{x\}) = \varnothing.$$

显然, $\mathscr{M} = \{U_x \mid x \in X\}$ 是 X 的一个开覆盖. 由 X 是紧致空间, $\mathscr{M}$ 有一个有限子覆盖

$$\{U_{x_1}, U_{x_2}, \cdots, U_{x_n}\}.$$

但每个 U_{x_i} 至多包含 A 的一个点,这与 A 为无穷点集矛盾. 从而, A 有聚点. X 是列紧空间.

我们知道,度量空间(S, ρ)是列紧度量空间的充分必要条件是(S, ρ)是紧致空间(见本篇第一章第四节定理 9),在一般拓扑空间中,列紧空间不必是紧致空间.

例 8 **设 X 是正整数集, $\forall$ 正整数 n,集合**

$$A_n = \{1, 2, \cdots, n\}$$

叫做正整数集的一个片断. 记

$$\mathscr{T} = \{X, \varnothing, A_1, A_2, \cdots, A_n, \cdots\},$$

则易知$(X, \mathscr{T})$是一个拓扑空间.

$$\mathscr{M} = \{A_1, A_2, \cdots, A_n, \cdots\}$$

是 X 的一个开覆盖，但显然没有有限子覆盖. 因此，X 不是紧致空间.

另一方面，设 A 是 X 的一个无穷点集. 设 n_0 是 A 所包含的正整数之最小者. 那么，当 $n > n_0$ 时，$n_0 \in (A_n - \{n\})$，因而，n 是 A 的聚点. 而 A 是无穷点集，在大于 n_0 的所有正整数中必有属于 A 的. 这样便知 X 是列紧空间.

定理 4　设拓扑空间$(X, \mathscr{T})$是列紧空间，T_1 空间，$\{F_n\}$是 X 的非空闭集序列，且

$$F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

那么，$\bigcap_{n=1}^{\infty} F_n \neq \varnothing$.

证　从每个 F_n 中取出一点 x_n. 记

$$E_1 = \{x \mid x \in X, x = \text{某个 } x_n\}.$$

若 E_1 是有限集，则在选出的点 x_n 中有无穷多个是相同的. 不妨记之为 a. 则 $a \in \bigcap_{n=1}^{\infty} F_n$. 结论成立.

若 E_1 是无限集. 因 F_1 是列紧空间 X 的闭集，故 F_1 是列紧集. E_1 是 F_1 的无穷子集，故 E_1 有聚点，且 $E_1' \subseteq F_1$. $\forall$ 正整数 n，记

$$E_n = \{x \mid x \in X, x = \text{某个 } x_k, k \geqslant n\}$$

X 是 T_1 空间，$E_1 - E_n$ 是有限集，因此，$E_n' = E_1'$. 于是

$$E_n' \subseteq F_n' \subseteq F_n,$$

即有 $E_1' \subseteq$ 每个 F_n，故 $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$.

定理 4 显然是数学分析中区间套定理的推广，也叫做Cantor 定理.

定理 5　设拓扑空间$(X, \mathscr{T})$是列紧空间，T_1 空间，则 X 的每个可数开覆盖$\{U_n\}$都有一个有限子覆盖.

证　不妨假设 U_n 是 X 的可数无穷覆盖. 记

$$V_n = \bigcup_{k=1}^{n} U_n, \ F_n = X - V_n = V_n^c$$

则 $V_n \in \mathscr{T}$, F_n 是闭集且 $F_{n+1} \subseteq F_n$.

假如每个 $F_n \neq \varnothing$，则由 Cantor 定理，$F = \bigcap_{n=1}^{\infty} F_n \neq \varnothing$. 于是

$$\bigcup_{k=1}^{n} U_n = \bigcup_{k=1}^{n} V_n = \bigcup_{k=1}^{n} (X - F_n) = X - \bigcap_{n=1}^{\infty} F_n \neq X,$$

这与$\{U_n\}$覆盖 X 矛盾. 因此, ∃ 正整数 n, 使得 $F_n = \varnothing$. 于是, $\{U_1, U_2, \cdots, U_n\}$就是一个所求的有限子覆盖.

三、覆盖式列紧空间

定义 3 每一个可数覆盖都有一个有限子覆盖的拓扑空间叫做覆盖式列紧空间.

拓扑空间的子集如果作为子空间是覆盖式列紧空间, 叫做拓扑空间的覆盖式列紧集.

定理 6 覆盖式列紧空间是列紧空间.

证 设覆盖式列紧空间$(X, \mathscr{T})$, 假设 X 不是列紧空间. 则 X 有一可数无穷点集

$$A = \{x_1, x_2, \cdots, x_n, \cdots\}$$

无聚点, 因而 A 是闭集, $A^c = X - A$ 是开集.

$\forall x_k \in A$, $\exists U_k \in \mathscr{T}$, 使得 $x_k \in U_k \cap (A - \{x_k\}) = \varnothing$. 于是 X 的可数开覆盖

$$\{A^c, U_1, U_2, \cdots, U_n, \cdots\}$$

没有有限子覆盖, 矛盾.

由定理 5, 可以看到, 每个列紧 T_1 空间是覆盖式列紧空间. 再由定理 6, 便知**对于 T_1 空间, 覆盖式列紧性与列紧性等价.**

显然, 紧致空间是覆盖式列紧空间. 反之, 则不然. 反例可参阅书末的文献.

设拓扑空间$(X, \mathscr{T})$的任意非空闭集序列 $\{F_n\}$, $F_{n+1} \subseteq F_n$, 都有

$$\bigcap_{n=1}^{\infty} F_n \neq \varnothing.$$

这一条件, 叫做Cantor 条件.

定理 7 拓扑空间$(X, \mathscr{T})$是覆盖式列紧空间⇔X 满足 Cantor 条件.

证 必要性: 假定有 X 的非空闭集序列 $\{F_n\}$, $F_{n+1} \subseteq F_n$, 而 $\bigcap_{n=1}^{\infty} F_n = \varnothing$. 令 $U_n = X - F_n$, 则 $\bigcup_{n=1}^{\infty} U_n = X$, 因此$\{U_n\}$ 是 X 的一个可数开覆盖. 于是有一个有限子覆盖

$$\{U_{n_1}, U_{n_2}, \cdots, U_{n_k}\} \ (n_1 < n_2 < \cdots < n_k).$$

于是 $\bigcup_{i=1}^{k} U_{n_i} = U_{n_k}$, 从而 $U_{n_k} = X$. 那么 $F_{n_k} = \varnothing$. 矛盾.

充分性: 由定理 5、定理 4 即知.

四、序列式列紧空间

定义 4 空间中的每个序列都有一个收敛子列的拓扑空间叫做序列式列紧空间.

拓扑空间的子集如果作为子空间是序列式列紧空间，叫做拓扑空间的序列式列紧集.

定理 8 序列式列紧空间是覆盖式列紧空间.

证 设$\{F_n\}$是序列式列紧空间$(X, \mathscr{T})$中的非空闭集序列，且$F_{n+1} \subseteq F_n$. 在每个F_n中选出一点x_n，则得X的序列$\{x_n\}$，$\{x_n\}$有收敛子列$\{x_{n_k}\}$，设其极限为x. 于是，$x \in \overline{F}_n = F_n$，故$x \in \bigcap_{n=1}^{\infty} F_n$. 由定理 7 便知$X$是覆盖式列紧空间.

由上面的论述，可以看到，紧致性强于覆盖式列紧性，覆盖式列紧性强于列紧性. 序列式列紧性强于覆盖式列紧性，但可以证明，紧致性并非强于序列式列紧性，序列式列紧性也并非强于紧致性.

定理 9 满足公理A_1的列紧T_1空间，是序列式列紧空间.

证 设拓扑空间$(X, \mathscr{T})$满足公理A_1，是列紧空间，T_1空间. 设$\{x_n\}$是X中的序列. 用A表示这个序列的点的集合. 因X是列紧空间. 故$\exists x \in X$，使得x的任意邻域V，包含$\{x_n\}$的无穷多项.

因X满足公理A_1，故点x处有可数邻域基$\{W_n\}$，$W_{n+1} \subseteq W_n$. 取$x_{n_k} \in W_k$ $(n_{k+1} > n_k,\ k = 1, 2, \cdots)$，则得$\{x_n\}$的收序子列$\{x_{n_k}\}$，其极限为$x$.

这便证明了X是序列式列紧空间.

由定理 9，立得

定理 10 满足公理A_1的紧致T_1空间是序列式列紧空间.

综前述可知

定理 11 对于满足公理A_1的T_1空间，

$$\text{列紧性} \Leftrightarrow \text{序列式列紧性} \Leftrightarrow \text{覆盖式列紧性}.$$

定理 12 满足公理A_2的覆盖式列紧空间是紧致空间.

证 设覆盖式列紧空间$(X, \mathscr{T})$满足公理A_2. 设$\mathscr{M}$是X的一个开覆盖. 由 Lindelof 定理，$\mathscr{M}$有一个可数子覆盖$\mathscr{U} \subseteq \mathscr{M}$. 再由$X$是覆盖式列紧空间. $\mathscr{U}$有一个有限子覆盖$\mathscr{M}_1 \subseteq \mathscr{U}$. $\mathscr{M}_1$当然也是$\mathscr{M}$的有限子覆盖.

综前述，便有

定理 13 对于满足公理A_2的拓扑空间，

$$\text{紧致性} \Leftrightarrow \text{覆盖式列紧性}.$$

定理 14 对于满足公理A_2的T_1空间，

$$\text{紧致性} \Leftrightarrow \text{覆盖式列紧性} \Leftrightarrow \text{列紧性} \Leftrightarrow \text{序列式列紧性}.$$

习 题 2-2

1. 证明：非空集合 X 上的拓扑 $\mathscr{T}_\alpha$ 的交集 $\mathscr{T}=\bigcap_\alpha \mathscr{T}_\alpha$ 是 X 上的拓扑.

2. 设非空集合 $X=\{a, b, c\}$，试写出 X 上的所有的拓扑. 在这些拓扑中，是否存在 $\mathscr{T}_1$，$\mathscr{T}_2$，使得 $\mathscr{T}_1\cup\mathscr{T}_2$ 不是 X 的拓扑？

3. 设集合 $X=\{a, b, c\}$，$\mathscr{T}=\{X, \varnothing, \{a, b\}, \{c\}\}$. 指出拓扑空间$(X, \mathscr{T})$的所有的闭集与完备集.

4. F 是拓扑空间$(X, \mathscr{T})$的闭集，且 $F'\subseteq E\subseteq F$. 证明：E 是闭集.

5. F 是拓扑空间$(X, \mathscr{T})$的自密集，且 $F\subseteq E\subseteq F'$. 证明：E 是自密集.

6. 证明：自密集的导集与闭包都是自密集.

7. E 是拓扑空间$(X, \mathscr{T})$的子集. 证明：E 无处稠密 $\Leftrightarrow \mathrm{Int}\overline{E}=\varnothing$.

8. E 是拓扑空间$(X, \mathscr{T})$的子集. 证明：E 是闭集且无处稠密 $\Leftrightarrow \exists U\in\mathscr{T}$，使得 $E=\partial U$.

9. 设拓扑空间 X 与 Y，$x_0\in X$. 映射 $f:X\to Y$ 在点 x_0 连续. X 中的序列$\{x_n\}$收敛于 x_0. 证明：Y 中的序列$\{f(x_n)\}$收敛于 $f(x_0)$. 举例说明这个命题的逆命题不真.

10. 证明：在拓扑空间 X 中，X 的开子集 G 的开子集 A 是 X 的开子集；X 的闭子集 F 的闭子集 B 是 X 的闭子集.

11. 设$(A, \mathscr{T}_A)$是拓扑空间$(X, \mathscr{T})$的一个子空间. B 是 A 的子集. B_A'，$\overline{B}_A$，Int_AB，Ext_AB，∂_AB 分别表示 B 在子空间$(A, \mathscr{T}_A)$的导集，闭包，内部，外部和边界. 指出下列等式是否正确，若正确，则加以证明；若否，则举出反例.

(1) $\overline{B}_A=\overline{B}$； (2) $\overline{B}_A=\overline{B}\cap A$；

(3) $B_A'=B'$； (4) $B_A'=B'\cap A$；

(5) $\mathrm{Int}_AB=\mathrm{Int}B$； (6) $\mathrm{Int}B\subseteq\mathrm{Int}_AB$；

(7) $\mathrm{Ext}_AB=\mathrm{Ext}B$； (8) $\mathrm{Ext}_AB=\mathrm{Ext}B\cap A$；

(9) $\partial_AB=\partial B$； (10) $\partial_AB=\partial B\cap A$；

(11) $\mathrm{Int}_AB=\mathrm{Int}B\cap A$.

12. 证明：直线 $\mathbf{E}^1$ 的非空连通集必是至少包含两个点的区间.

13. 证明：直线 $\mathbf{E}^1$ 挖去一个点 a 后是非连通的，而平面 $\mathbf{E}^2$ 挖去一个点后是连通的.

14. 证明：拓扑空间 X 非连通$\Leftrightarrow$存在连续映射 $f:X\to\mathbf{E}^1$，使得 $f(X)$只含 2 个元素.

15. 证明：如果连通的度量空间(S, ρ)至少包含 2 个元素，则存在连续映射 $f:X\to[0, 1]$，使得 $f(X)=[0, 1]$.

16. 证明：本章第二节例 6 中的 $\overline{A}$ 不是道路连通集.

17. 道路连通集的闭包是否也是道路连通集？

18. 拓扑空间 X 的连通分支一定是 X 的开集吗？

19. 连通空间的子空间一定是连通空间吗？

20. 两个连通空间的积空间是连通空间吗？

21. 举例说明，T_1 空间的收敛序列可能极限不惟一.

22. 证明：T_0 空间的子空间与积空间都是 T_0 空间.

23. 证明：T_1 空间的子空间与积空间都是 T_1 空间.

24. 设 T_1 空间的点集 A 是闭集，至少包含 2 个点，并且是连通的. 证明：A 是完备集.

25. T_2 空间的子空间与积空间是否是 T_2 空间？

26. 设 A 是拓扑空间$(X, \mathscr{T})$的非空点集，Y 是 T_2 空间. 连续映射 $f: X \to Y$. 在 $x_0 \in \overline{A}$ 处有极限. 证明：f 在 x_0 处的极限是惟一的.

27. T_3 空间的子空间与积空间是否是 T_3 空间？

28. 正规空间的子空间与积空间是否是正规空间？

29. T_4 空间的子空间与积空间是否是 T_4 空间？

30. 设拓扑空间$(X, \mathscr{T})$满足第一可数性公理，E 是 X 的非空点集.

证明：$x \in \overline{E} \Leftrightarrow E$ 中有收敛于 x 的序列$\{x_n\}$.

31. 设拓扑空间$(X, \mathscr{T})$满足第一可数性公理，X 中的收敛序列有惟一的极限.

证明：$(X, \mathscr{T})$是 Hausdorff 空间.

32. 证明：度量空间是可分空间的充分必要条件是满足第二可数性公理.

33. 证明：紧致性、列紧性、覆盖式列紧列、序列式列紧性都是拓扑性质.

34. 举例说明紧致 T_1 空间的紧致集不必是闭集.

35. 证明：T_2 空间的两个紧致集的交集是紧致集.

36. 证明：正则空间的紧致集的闭包是紧致集.

参 考 文 献

1. 熊金城. 点集拓扑学讲义. 北京:人民教育出版社,1981
2. M. A. Armstrong. 基础拓扑学. 孙以丰,译. 北京:北京大学出版社,1983
3. J. L. Kelley. 一般拓扑学. 吴从炘,吴让泉译. 北京:科学出版社,1982
4. 李孝传,陈玉清. 一般拓扑学. 北京:人民教育出版社. 1982
5. 江泽涵. 拓扑学引论. 上海:上海科学技术出版社,1978
6. G. L. Naber. Topological Methods in Euclidean Spaces. Cambridge, 1980.

第三篇　泛 函 分 析

泛函分析是研究抽象空间及其在空间上定义的算子(映射)性质的一门数学学科.泛函分析中,常把一些分析的、代数的、几何的对象(函数、数列、矩阵、算子等)作为某种特定意义下"空间"的元素,再在抽象"空间"中加以一般性的讨论,找出其共同规律性的东西.这里所谓"空间"是指满足一定公理的元素所组成的基本集合.

第一章　Banach空间及其上的算子

上一篇研究了一类特殊的拓扑空间——度量空间.在应用中常见的抽象空间不但具有拓扑结构,而且具有一定的代数结构.例如,考察物理系统的状态所构成的集合,人们不但要研究其收敛性等拓扑性质,而且总希望状态允许有加法和与数的乘法运算,并使其代数结构与拓扑结构有机地结合起来成为性质优良的线性拓扑空间.

本章着重研究应用中最重要的一类特殊的线性拓扑空间——赋范线性空间与完备的赋范线性空间及其上定义的线性算子与线性泛函的理论.

第一节　赋范线性空间与Banach空间

一、线性空间

定义1　设 X 是一非空集合,K 是数域(实数域 $\mathbf{R}$ 或复数域 $\mathbf{C}$).如果在 X 中定义了元素的加法和数与元素的乘法,且满足下列条件:

(A) $\forall x, y \in X$, $\exists u \in X$,叫做 x, y 的和,记为 $u = x + y$,且满足

(1) $x + y = y + x$;

(2) $(x + y) + z = x + (y + z)$;

(3) 存在零元素 θ, $\forall x \in X$ 有　$x + \theta = x$;

(4) $\forall x \in X$, $\exists(-x) \in X$ 有 $\quad x+(-x)=\theta$.

(B) $\forall x \in X$ 及 $\forall \lambda \in K$, $\exists u \in X$, 叫做 λ 与 x 的数乘, 记为 $u=\lambda x$, 且 $\forall x, y \in X, \lambda, \mu \in K$ 满足

(5) $1x=x$;

(6) $\lambda(\mu x)=(\lambda\mu)x$;

(7) $\lambda(x+y)=\lambda x+\lambda y$;

(8) $(\lambda+\mu)x=\lambda x+\mu x$.

则说 X 是数域 K 上的线性空间或向量空间. 其中的元素叫做向量或点.

零向量 θ 有时简记为 0, 它未必是数 0.

特别地, 实数域 $\mathbf{R}$ 上的向量空间叫做实空间; 复数域 $\mathbf{C}$ 上的向量空间叫做复空间.

由定义不难验证, 对 $\lambda \in K, x \in X$ 恒有

$$0x=\theta;\ \lambda\theta=\theta;\ (-1)x=-x.$$

例 1 数列空间 $X=\{x=\{\xi_i\}_{i=1}^m \mid \xi_i \in K$, 且 x 满足性质 $P\}$, 这里 m 为有限或 ∞. 定义

$$x+y=\{\xi_i+\eta_i\}_{i=1}^m,$$
$$\lambda x=\{\lambda\xi_i\}_{i=1}^m,$$

其中 $x=\{\xi_i\}_{i=1}^m$, $y=\{\eta_i\}_{i=1}^m \in X$, $\lambda \in K$, 且恒有 $x+y, \lambda x \in X$. 则易验证 X 成为线性空间.

例 2 函数空间 $\quad X=\{f \mid f: \Omega \to K$ 且满足性质 $P\}$. 定义

$$(f+g)(x)=f(x)+g(x),$$
$$(\lambda f)(x)=\lambda f(x),$$

其中 $f, g \in X$, $\lambda \in K$, 且恒有 $f+g, \lambda f \in X$, 则易验证 X 成为线性空间.

后面常用的有关线性空间中的重要术语、记号与概念, 现简单罗列如下:

设 X, Y 都为 K 上的线性空间.

(1) 子空间

若 $X_0 \subseteq X$, 且 $\forall x, y \in X_0$, $\lambda \in K$, 恒有 $x+y, \lambda x \in X_0$, 则易证 X_0 为线性空间, 说 X_0 是 X 的线性子空间(简称子空间). 当 $\{\theta\} \neq X_0 \subset X$ 时, 说 X_0 是 X 的真子空间.

(2) 张成空间

若 $M \subseteq X$, 易证子集 $\left\{\sum_{i=1}^n a_i x_i \mid x_i \in M, a_i \in K, n \text{ 有限}\right\}$ 是 X 的子空间. 该子空间叫做 M 的张成空间, 记为 span M, 简称 M 的和.

易证 span M 为包含 M 的最小的子空间.

例 3 在 $\mathbf{E}^3$ 中,取 $M = \{e_1, e_2\}$,其中 $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$,则 span $M = xOy$ 平面,为 $\mathbf{E}^3$ 的真子空间.

(3) **基与维数**

若$\{e_1, e_2, \cdots, e_n\}$是 X 中线性无关的向量组,$X = \text{span}\{e_1, e_2, \cdots, e_n\}$,则说$\{e_1, e_2, \cdots, e_n\}$是 X 的一组基.基中所含向量的个数 n 叫做空间 X 的维数,记为 $\dim X$.

特别地,当 $X = \{\theta\}$,说 X 是零维空间.

若 $\dim X < +\infty$ 时,说 X 是有限维空间,否则说 X 是无限维空间.

例 4 $\mathbf{E}^n$ 是有限维空间,l^p, $C[a, b]$, $L^p[a, b]$都是无限维空间.

(4) **线性算子与线性泛函**

若 $A:X\to Y$ 满足:$\forall x_1, x_2 \in X$, $\alpha \in K$ 有

(1) $A(x_1 + x_2) = Ax_1 + Ax_2$ (可加性);

(2) $A(\alpha x_1) = \alpha Ax_1$ (齐次性),

则说 A 是 X 到 Y 中的线性算子.

注意 对线性算子而言,习惯上总把象$A(x)$简记为 Ax.

特别地,当 $Y = K$ 时,线性算子$f:X\to K$叫做线性泛函,且当 $K = \mathbf{R}$ 或 $\mathbf{C}$ 时,相应叫做实线性泛函或复线性泛函.

例 5 定义 $Af = \int_a^x f(t)\mathrm{d}t$, $f \in C[a, b]$,易验证

$$A:C[a, b] \to C[a, b]$$

是线性算子.

例 6 定义 $Df = \frac{\mathrm{d}}{\mathrm{d}x}f(x)$, $f \in C^{(1)}[a, b]$,易验证

$$D:C^{(1)}[a, b]\to C[a, b]$$

是线性算子.

例 7 若 $f(x) = \int_a^b x(t)\mathrm{d}t$, $x \in C[a, b]$,则 $f:C[a, b]\to K$ 是线性泛函.但是 $g(x) = \int_a^b \sqrt{1 + (x'(t))^2}\mathrm{d}t$, $x \in C^{(1)}[a, b]$,则易知泛函

$$g:C^{(1)}[a, b]\to \mathbf{R}$$

不是线性的.

(5) **同构**

若存在可逆的线性算子 $A:X\to Y$,则说 A 是 X 到 Y 上的线性同构映射,简

称为同构映射，并说 X 与 Y 是同构的.

例 8 n 维实(或复)线性空间 X 总与 $\mathbf{E}^n$(或 $\mathbf{C}^n$)同构.

注意 同构映射 $A:X\to Y$ 的逆映射 $A^{-1}:Y\to X$ 也是同构映射. 从代数性质而言，同构的两空间 X, Y 可视为相同的空间.

有限维空间是线性代数的主要研究对象，泛函分析研究的主要对象是无限维空间，它不但要研究其代数性质，还要研究与代数结构有机地相结合的拓扑性质，这就是线性拓扑空间理论.

定义 2 **设 X 是线性空间，且 X 有 T_2 拓扑，使对此拓扑而言，X 上的线性运算都是连续的，则说 X 是线性拓扑空间.**

定义 3 **设 M 是线性空间 X 的一子集，若 $\forall x,\ y\in M$ 为端点的线段 $\{\alpha x+(1-\alpha)y \mid \alpha\in[0,1]\}\subseteq M$，则说 M 是 X 的凸子集.**

线性空间 X 及其任一子空间都是凸集.

二、赋范线性空间与 Banach 空间

定义 4 **设数域 K 上的线性空间 X 上定义有实泛函 $F:X\to\mathbf{R}$，记**

$$F(x)=\|x\|,$$

若 $\forall x,\ y\in X,\ \alpha\in K$ 满足：

1°正定性： $\|x\|\geqslant 0$，**且** $\|x\|=0\Leftrightarrow x=\theta$；

2°正齐性： $\|\alpha x\|=|\alpha|\,\|x\|$；

3°三角不等式： $\|x+y\|\leqslant\|x\|+\|y\|$，

则说 $\|x\|$ 是 x 的范数，$(X,\ \|\cdot\|)$ 叫做赋范线性空间，简记为 X. 通常称定义中的条件 1°, 2°, 3°为范数公理.

在赋范线性空间 X 中可以定义距离

$$d(x,\ y)=\|x-y\|,\ \forall x,\ y\in X. \tag{1}$$

易知 $\forall x,\ y,\ z\in X$，满足：

(1) $d(x,\ y)=\|x-y\|\geqslant 0$ 且

$d(x,\ y)=0\Leftrightarrow\|x-y\|=0\Leftrightarrow x=y$；

(2) $d(x,\ y)=\|x-y\|=\|y-x\|=d(y,\ x)$；

(3) $d(x,\ y)=\|x-y\|=\|x-z+z-y\|$

$\leqslant\|x-z\|+\|z-y\|=d(x,\ z)+d(z,\ y)$.

d 叫做由范数 $\|\cdot\|$ 决定的距离. 今后如不特殊声明，赋范线性空间 $(X,\|\cdot\|)$ 总是指按式(1)中距离 d 而成为度量空间. 所以，度量空间中的所有结果都可应用于赋范线性空间.

若$(X,\ d)$是线性的度量空间，那么 d 是否是 X 上某范数决定的呢？即

$\|x\| = d(x, \theta)$ 是否满足范数公理呢？请读者自己思考.

例 9 设数列空间 $S = \{\{\xi_i\}_{i=1}^{\infty} \mid \xi_i \in \mathbf{C}\}$ 中定义

$$d(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|\xi_i - \eta_i|}{1+|\xi_i - \eta_i|},$$

其中 $x = \{\xi_i\}$，$y = \{\eta_i\}$. 易验证(S, d)是线性度量空间，但 $\|x\| = d(x, \theta) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|\xi_i|}{1+|\xi_i|}$ 不满足范数公理 2°. 因为当 $\alpha \neq 0, 1$ 时 $\|\alpha x\| \neq |\alpha| \|x\|$ $(\forall x \neq \theta)$.

因此，线性度量空间要成为赋范线性空间是有条件的.

定理 1 **设 (X, d) 是线性度量空间，则 (X, d) 是由赋范线性空间 $(X, \|\cdot\|)$ 决定的充分必要条件是 d 满足如下两条件：**

(1) $d(x, y) = d(x-y, \theta)$, $\quad \forall x, y \in X$;

(2) $d(\alpha x, \theta) = |\alpha| d(x, \theta)$, $\quad \forall x \in X, \alpha \in K$.

证 必要性：因为 $d(x, y) = \|x - y - \theta\| = d(x-y, \theta)$,

$$d(\alpha x, \theta) = \|\alpha x - \theta\| = \|\alpha x\| = |\alpha| \|x\| = |\alpha| d(x, \theta).$$

充分性：令 $\|x\| = d(x, \theta)$,

(1) $\|x\| = d(x, \theta) \geqslant 0$ 且 $\|x\| = d(x, \theta) = 0 \Leftrightarrow x = \theta$;

(2) $\|\alpha x\| = d(\alpha x, \theta) = |\alpha| d(x, \theta) = |\alpha| \|x\|$;

(3) $\|x+y\| = d(x+y, \theta) = d(x-(-y), \theta) = d(x, -y)$

$\leqslant d(x, \theta) + d(\theta, -y) = \|x\| + \|-y\| = \|x\| + \|y\|$,

故 $\|x\| = d(x, \theta)$ 是 X 上的范数.

三、赋范线性空间中的极限

关于赋范线性空间中点列的收敛性及相关概念，只要在由范数导出的距离 $d(x, y) = \|x - y\|$ 之下来讨论，就可得到相关定义和结果.

定义 5 **设 $x_n, x \in (X, \|\cdot\|)$ $(n=1, 2, 3, \cdots)$,**

(1) 若$\lim\limits_{n\to\infty} \|x_n - x\| = 0$，则说$\{x_n\}$依范数收敛于 x，或强收敛于 x，简称收敛于 x，记为

$$\lim_{n\to\infty} x_n = x \quad \text{或} \quad x_n \to x (n \to \infty);$$

(2) 若 $\forall \varepsilon > 0$，存在 N，使得当 $m, n > N$ 时都有

$$\|x_m - x_n\| < \varepsilon,$$

则说$\{x_n\}$是$(X, \|\cdot\|)$中的Cauchy 列(或基本列).

定理 2 设$(X, \|\cdot\|)$是赋范线性空间，则

(1) $\|x\|$是x的连续泛函；

(2) 若$x_n, x \in X$, $x_n \to x$，则$\{\|x_n\|\}$有界；

(3) 线性运算是连续的，即若$x_n, x, y_n, y \in X$, $\lambda_n, \lambda \in K$，且

$$x_n \to x,\ y_n \to y,\ \lambda_n \to \lambda,$$

则

$$\lambda_n x_n \to \lambda x,\ x_n + y_n \to x + y.$$

证 (1) $\|\cdot\|:(X, \|\cdot\|) \to \mathbf{R}$，若$x_n \to x$，则$\|x_n - x\| = d(x_n, x) \to 0$. 由距离函数$d(x, y)$的连续性知$d(x_n, \theta) \to d(x, \theta)$，即$\|x_n\| \to \|x\|$，所以$\|x\|$是$x$的连续泛函.

(2) 由(1)知$\{x_n\}$收敛于x，则$\{\|x_n\|\}$收敛于$\|x\|$，故$\|x_n\|$有界.

(3) 由 $\|\lambda_n x_n - \lambda x\| = \|(\lambda_n - \lambda)x_n + \lambda(x_n - x)\|$

$$\leqslant |\lambda_n - \lambda|\ \|x_n\| + |\lambda|\ \|x_n - x\|$$

与$\|(x_n + y_n) - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\|$ 及(2)，知线性运算都是连续的.

四、Banach(巴拿赫)空间

定义 6 设X是一赋范线性空间，如果X按照距离$d(x, y) = \|x - y\|$是完备的，则说X是Banach 空间.

例 10 在下列线性空间中分别定义范数：

(1) $\mathbf{E}^n$: $\|x\| = \left(\sum_{i=1}^{n} |\xi_i|^2\right)^{\frac{1}{2}}$, $x = (\xi_1, \xi_2, \cdots, \xi_n) \in \mathbf{E}^n$;

(2) $l^p(1 \leqslant p < \infty$: $\|x\| = \left(\sum_{i=1}^{\infty} |\xi_i|^p\right)^{\frac{1}{p}}$, $x = (\xi_1, \xi_2, \cdots, \xi_n, \cdots) \in l^p$;

(3) l^∞: $\|x\| = \sup_i |\xi_i|$, $x = (\xi_1, \xi_2, \cdots, \xi_n, \cdots) \in l^\infty$;

(4) $C[a, b]$: $\|x\| = \max_{a \leqslant t \leqslant b} |x(t)|$, $x \in C[a, b]$;

(5) $L^p[a, b]$: $\|x\| = \left(\int_a^b |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}}$, $x \in L^p[a, b]$;

(6) $L^\infty[a, b]$: $\|x\|_\infty = \inf_{m(A)=0} \sup_{t \in [a, b]-A} |x(t)|$, $x \in L^\infty[a, b]$.

易证它们都满足范数公理，故相应的空间都成为赋范线性空间，并且由于它们的范数所决定的度量空间都是完备的. 因此，它们都是 Banach 空间.

今后若不加说明，仍分别记为$\mathbf{E}^n$, l^p, l^∞, $C[a, b]$, $L^p[a, b]$, $L^\infty[a, b]$.

例 11 若在$C[a, b]$中定义$\|x\|_1 = \int_a^b |x(t)|\,\mathrm{d}t$，则$(C[a, b], \|\cdot\|_1)$

也成为赋范线性空间，但由 $\|\cdot\|_1$ 决定的度量

$$d_1(x, y) = \|x-y\|_1 = \int_a^b |x(t)-y(t)| \mathrm{d}t$$

构成不完备的度量空间 $(C[a, b], d_1)$，故 $(C[a, b]), \|\cdot\|_1)$ 不是 Banach 空间.

在赋范线性空间中，由于范数刻画了向量的长度，因此，赋范空间中的概念具有更强的几何直观性.

例 12 证明：赋范线性空间 X 中的球均是凸集.

证 设 $V(x_0, r)$ 为 X 中以 x_0 为中心，$r>0$ 为半径的开球. $\forall x, y \in V(x_0, r)$. 令 $z = \alpha x + (1-\alpha)y$, $0 \leqslant \alpha \leqslant 1$ 则有

$$\begin{aligned}\|z-x_0\| &= \|\alpha x + (1-\alpha)y - [\alpha x_0 + (1-\alpha)x_0]\| \\ &\leqslant \alpha\|x-x_0\| + (1-\alpha)\|y-x_0\| \\ &< \alpha r + (1-\alpha)r = r,\end{aligned}$$

即 $z \in V(x_0, r)$，故 $\{\alpha x + (1-\alpha)y : 0 \leqslant \alpha \leqslant 1\} \subseteq V(x_0, r)$.

因此，$V(x_0, r)$ 为凸集.

类似可证闭球 $V[x_0, r]$ 也为凸集.

五、赋范线性空间中的级数

由于赋范线性空间不但有拓扑结构，而且有线性的代数结构，所以与一般的度量空间不同，在赋范空间 $(X, \|\cdot\|)$ 中可以引进无穷级数的概念.

定义 7 设 $\{x_n\} \subseteq (X, \|\cdot\|)$，无穷和式

$$x_1 + x_2 + \cdots + x_n + \cdots = \sum_{n=1}^{\infty} x_n$$

叫做无穷级数，简称为级数.

$S_N = \sum_{n=1}^{N} x_n$ 叫做级数 $\sum_{n=1}^{\infty} x_n$ 的前 N 项部分和.

若存在 $x_0 \in X$，使得 $\lim_{N\to\infty} S_N = x_0$，则说级数 $\sum_{n=1}^{\infty} x_n$ 收敛于 x_0，且 x_0 叫做级数 $\sum_{n=1}^{\infty} x_n$ 的和.

若 $\{S_N\}$ 在 X 中不收敛，则说级数 $\sum_{n=1}^{\infty} x_n$ 发散.

定理 3 设 X 为 Banach 空间，$\{x_n\} \subseteq X$，则级数 $\sum_{n=1}^{\infty} x_n$ 收敛的充分必要条件

是 $\forall \varepsilon > 0$，$\exists N$，使得当 $m > n > N$ 时恒有 $\left\|\sum_{i=n}^{m} x_i\right\| < \varepsilon$ 成立.

证 必要性：设 $\lim_{n\to\infty}\sum_{i=1}^{n} x_i = x_0$，故 $\forall \varepsilon > 0$，存在 N，使得当 $m > n > N$ 时恒有

$$\left\|\sum_{i=1}^{m} x_i - x_0\right\| < \frac{\varepsilon}{2}, \left\|\sum_{i=1}^{n-1} x_i - x_0\right\| < \frac{\varepsilon}{2},$$

从而有

$$\begin{aligned}\left\|\sum_{i=n}^{m} x_i\right\| &= \left\|\sum_{i=1}^{m} x_i - x_0 + x_0 - \sum_{i=1}^{n-1} x_i\right\| \\ &\leqslant \left\|\sum_{i=1}^{m} x_i - x_0\right\| + \left\|\sum_{i=1}^{n-1} x_i - x_0\right\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

充分性：因 $\forall \varepsilon > 0$，存在 N，使得当 $m > n > N$ 时恒有 $\left\|\sum_{i=n}^{m} x_i\right\| < \varepsilon$，即 $\{x_n\}$ 是 Banach 空间 X 中的 Cauchy 列，故级数 $\sum_{n=1}^{\infty} x_n$ 收敛.

第二节　有界线性算子

一、有界线性算子（连续线性算子）

设 X，Y 都为数域 K 上的线性空间，$A: X \to Y$ 为线性算子，

$$N(A) = \{x \in X \mid Ax = \theta\}$$

叫做算子 A 的零空间，

$$R(A) = \{Ax \mid x \in X\} (\subseteq Y)$$

叫做算子 A 的值域.

定理 1　设 $A: X \to Y$ 为线性算子，则

(1) $A\theta = \theta$；

(2) $N(A)$ 为 X 中线性子空间；

(3) 若 A^{-1} 存在，则 $A^{-1}: Y \to X$ 也为线性算子；

(4) A^{-1} 存在的充分必要条件是 $N(A) = \{\theta\}$，且 $R(A) = Y$.

证　(1) $A\theta = A(0\theta) = 0A\theta = \theta$.

(2) 设 $x_1, x_2 \in N(A)$, $\alpha, \beta \in K$,必有

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2 = \alpha\theta + \beta\theta = \theta,$$

所以 $\alpha x_1 + \beta x_2 \in N(A)$, 即 $N(A)$为 X 的子空间.

(3) 若 A^{-1}存在, $\forall y_1, y_2 \in Y$, $\forall \alpha, \beta \in K$. $\exists x_1, x_2 \in X$,使得 $Ax_1 = y_1$, $Ax_2 = y_2$.

由于

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2 = \alpha y_1 + \beta y_2,$$

因此

$$A^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha A^{-1} y_1 + \beta A^{-1} y_2,$$

即 $A^{-1}: Y \to X$ 为线性算子.

(4) 必要性:若 A^{-1} 存在,显然$R(A)=Y$,且 A 是单射. 设 $x \in N(A)$, 则 $Ax = \theta = A\theta$, 故有 $x = \theta$, 即 $N(A) = \{\theta\}$.

充分性:若 $N(A) = \{\theta\}$, 则由 $Ax = 0$ 必有 $x = \theta$. 因此,若 $Ax_1 = Ax_2$,有 $A(x_1 - x_2) = Ax_1 - Ax_2 = \theta$,故$x_1 - x_2 = \theta$,所以$x_1 = x_2$, 即 A 为 X 到 Y 中的单射;又因 $R(A) = Y$, 故 A 又是满射,从而$A^{-1}: Y \to X$存在.

定义 1 设 X, Y 都为赋范线性空间,$A: X \to Y$ 为线性算子,若 $\exists M > 0$,使得

$$\| Ax \|_Y \leqslant M \| x \|_X, \quad \forall x \in X,$$

则说 A 是有界线性算子,否则说 A 是无界线性算子.

今后在不致误会情况下,$\| Ax \|_Y$, $\| x \|_X$ 分别简记为 $\| Ax \|$, $\| x \|$.

定理 2 线性算子 $A: X \to Y$ 为有界的充分必要条件是 A 将 X 中任一有界集 Ω 映成为 Y 中有界集$A(\Omega)$.

证 必要性:设 Ω 为 X 中有界集,即有 $K > 0$ 使得 $\| x \| \leqslant K$, $\forall x \in \Omega$. 因 A 是有界线性算子,故有常数 $M > 0$,使得

$$\| Ax \| \leqslant M \| x \| \leqslant MK, \quad \forall x \in \Omega,$$

即 $A(\Omega)$是 Y 中的有界集.

充分性:令 $S = \{x: \| x \| = 1, x \in X\}$ 为 X 中单位球面,因 S 有界,故存在 $M > 0$, 使得 $\| Ax \| \leqslant M$, $\forall x \in S$.

于是 $\forall x \in X$, $x \neq \theta$, 有 $\frac{x}{\| x \|} \in S$,从而有

$$\frac{1}{\| x \|} \| Ax \| = \| A \frac{x}{\| x \|} \| \leqslant M,$$

即 $\|Ax\| \leqslant M\|x\|$，$\forall x \neq \theta$，$x \in X$.

显然，$x=\theta$ 时上述不等式也成立.因此，A 是有界线性算子.

注意 从定理证明过程可以看出，线性算子若能将单位球面映成有界集，则它必为有界算子.

由于 X，Y 都是赋范线性空间时，它们都有由范数决定的度量拓扑，所以可以讨论线性算子 $A:X\to Y$ 的连续性.

定理 3 设 X，Y 都是同一数域 K 上的赋范线性空间，$A:X\to Y$ 是线性算子，则下列命题是等价的：

(1) A 是有界的；

(2) A 是一致连续的；

(3) A 是连续的；

(4) A 是在某点 $x_0 \in X$ 连续的.

证 (1)⇒(2)：因 A 有界，故存在 $M>0$，使得 $\forall x \in X$ 有 $\|Ax\| \leqslant M\|x\|$. 若 $\forall x_1, x_2 \in X$，则

$$\|Ax_1 - Ax_2\| = \|A(x_1 - x_2)\| \leqslant M\|x_1 - x_2\|.$$

故 A 是一致连续的.

(2)⇒(3)⇒(4) 显然.

(4)⇒(1)： 因 A 在 $x_0 \in X$ 连续，即 $x_n \to x_0$ 时有 $Ax_n \to Ax_0$，

① 证明 A 在 θ 连续：

若 $x_n \to \theta$，则 $x_n + x_0 \to x_0$.

因 $Ax_n + Ax_0 = A(x_n + x_0) \to Ax_0\ (n\to\infty)$，故 $Ax_n \to Ax_0 - Ax_0 = \theta = A\theta$.

从而 A 在 θ 连续.

② 证明 A 有界：

假设 A 无界，则 $\forall n$，$\exists x_n \in X$，使得

$$\|Ax_n\| > n\|x_n\| \neq 0. \tag{1}$$

令
$$y_n = \frac{x_n}{n\|x_n\|},$$

则
$$\lim_{n\to\infty}\|y_n\| = \lim_{n\to\infty}\frac{\|x_n\|}{n\|x_n\|} = 0, \quad 即 \quad y_n \to \theta,$$

因 A 在 θ 连续，故有 $Ay_n \to A\theta = \theta$. (2)

另一方面，由 A 的线性性和式(1)有

$$\|Ay_n\| = \left\|A\frac{x_n}{n\|x_n\|}\right\| = \frac{\|Ax_n\|}{n\|x_n\|} > \frac{n\|x_n\|}{n\|x_n\|} = 1.$$

这与式(2)式矛盾.

从而 A 有界.

二、有界线性算子空间

设 $A:X\to Y$ 是有界线性算子，则存在 $M>0$，使得 $\forall x\in X$ 有 $\|Ax\|\leqslant M\|x\|$. 因此 $\left\{\dfrac{\|Ax\|}{\|x\|}\mid x\neq\theta\right\}$ 是一个以 M 为上界的有界数集. 故它有有限的上确界.

定义 2 设 $A:X\to Y$ 是有界线性算子，$\|A\|=\sup\limits_{x\neq\theta}\dfrac{\|Ax\|}{\|x\|}$ 叫做 A 的范数.

显然 $\|A\|$ 是一个有限数.

定理 4 设 $A:X\to Y$，$B:Y\to Z$ 都是有界线性算子，则

(1) $\|Ax\|\leqslant\|A\|\,\|x\|,\quad\forall x\in X$;

(2) $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|=\sup\limits_{\|x\|\leqslant1}\|Ax\|$;

(3) $\|BA\|\leqslant\|B\|\,\|A\|$.

证 (1) 因 $\|A\|=\sup\limits_{x\neq\theta}\dfrac{\|Ax\|}{\|x\|}\geqslant\dfrac{\|Ax\|}{\|x\|},\quad\forall x\neq\theta$,

故 $\|Ax\|\leqslant\|A\|\,\|x\|$.

(2) 因
$$\begin{aligned}\|A\|&=\sup_{x\neq\theta}\frac{\|Ax\|}{\|x\|}=\sup_{x\neq\theta}\left\|A\frac{x}{\|x\|}\right\|=\sup_{\|y\|=1}\|Ay\|\\&\leqslant\sup_{\|y\|\leqslant1}\|Ay\|\leqslant\sup_{\|y\|\leqslant1}\|A\|\,\|y\|\\&\leqslant\sup_{\|y\|\leqslant1}\|A\|=\|A\|,\end{aligned}$$

所以 $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|=\sup\limits_{\|x\|\leqslant1}\|Ax\|$.

(3) $\forall x\in X$ 有 $\|BAx\|\leqslant\|B\|\,\|Ax\|\leqslant\|B\|\,\|A\|\,\|x\|$，故 $\dfrac{\|BAx\|}{\|x\|}\leqslant\|B\|\,\|A\|$， 从而

$$\|BA\|=\sup_{x\neq\theta}\frac{\|BAx\|}{\|x\|}\leqslant\|B\|\,\|A\|.$$

例 1 (1) 恒等算子 $I:X\to X,\quad Ix=x,\quad\forall x\in X$;

(2) 零算子 $O:X\to Y,\quad Ox=\theta,\quad\forall x\in X$,

则 I 和 O 都是线性算子，且

$$\|I\|=\sup_{x\neq\theta}\frac{\|Ix\|}{\|x\|}=\sup_{x\neq\theta}\frac{\|x\|}{\|x\|}=1,$$

$$\|O\|=\sup_{x\neq\theta}\frac{\|Ox\|}{\|x\|}=\sup_{x\neq\theta}\frac{0}{\|x\|}=0.$$

注意 由 $\|A\|=1$ 不能判定 $A=I$，但是，由 $\|A\|=0$ 可以判定 $A=\theta$.

例 2 设$\{e_1, e_2, \cdots, e_n\}$是 $\mathbf{E}^n$ 的一组基，$\{e'_1, e'_2, \cdots, e'_m\}$是 $\mathbf{E}^m$ 的一组基. 则 $A_1:\mathbf{E}^n\to\mathbf{E}^m$ 为线性算子的充分必要条件是 $\exists | A=(a_{ij})_{m\times n}$，使得

$$y=Ax=(a_{ij})_{m\times n}x,$$

其中 $x=(\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}\in\mathbf{E}^n$, $y=(\eta_1, \eta_2, \cdots, \eta_m)^{\mathrm{T}}\in\mathbf{E}^m$.

证 充分性：$\forall x\in\mathbf{E}^n$ 可惟一表为 $x=\sum\limits_{j=1}^{n}\xi_j e_j$, $\forall y\in\mathbf{E}^m$ 可惟一表成 $y=\sum\limits_{i=1}^{m}\eta_i e'_i$.

由

$$y=Ax=(a_{ij})_{m\times n}\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix} \tag{3}$$

记 $A_1:\mathbf{E}^n\to\mathbf{E}^m$，对应由式(3)决定.

显然，A_1 是线性算子.

必要性：$Ae_j\in\mathbf{E}^m(j=1, 2, \cdots, n)$，则 $Ae_j=\sum\limits_{i=1}^{m}a_{ij}e'_i(j=1, 2, \cdots, n)$，即

$$\begin{cases}Ae_1=a_{11}e'_1+a_{21}e'_2+\cdots+a_{m1}e'_m,\\ \cdots\cdots\cdots\cdots\\ Ae_n=a_{1n}e'_1+a_{2n}e'_2+\cdots+a_{mn}e'_m.\end{cases}$$

$$\text{故 } Ax=A\left(\sum_{j=1}^{n}\xi_j e_j\right)=\sum_{j=1}^{n}\xi_j Ae_j$$

$$=\sum_{j=1}^{n}\xi_j\sum_{i=1}^{m}a_{ij}e'_i=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}\xi_j\right)e'_i.$$

令 $\eta_i=\sum\limits_{j=1}^{n}a_{ij}\xi_j$, $i=1, 2, \cdots, m$.

则有

$$y=\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_m\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}=(a_{ij})_{m\times n}x.$$

注意 (1) 在有限维线性空间中，如选定基后，线性算子和矩阵是相对应的.

(2) 矩阵是线性算子的最直观模型.

(3) 设$(e_1, e_2, \cdots, e_n)$是 $\mathbf{E}^n$ 中一组基，基选定以后，则 $f:\mathbf{E}^n\to\mathbf{R}$是线性的充

分必要条件是 $\exists(a_1, a_2, \cdots, a_n)_{1\times n}$ 使得 $f(x)=\sum_{j=1}^{n}a_j\xi_j$，其中

$$x=(\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}\in \mathbf{E}^n.$$

例 3 在 $C^{(1)}[0, 1]$上定义范数 $\|x\|=\max\limits_{0\leqslant t\leqslant 1}|x(t)|$，则微分算子

$$D=\frac{\mathrm{d}}{\mathrm{d}t}:\ C^{(1)}[0, 1]\to C[0, 1]$$

是无界线性算子.

证 显然,D 是线性算子,下面利用定理 2 来证明 D 是无界的.

取 $x_n=t^n$,有 $\|x_n\|=\max\limits_{0\leqslant t\leqslant 1}|x_n(t)|=\max\limits_{0\leqslant t\leqslant 1}|t^n|=1$. 但是

$$\|Dx_n\|=\max_{0\leqslant t\leqslant 1}|x_n'(t)|=\max_{0\leqslant t\leqslant 1}|nt^{n-1}|=n.$$

因为 $\{x_n\}_{n=1}^{\infty}$ 是 $C^{(1)}[0, 1]$ 中的有界集,而 $\{Dx_n\}_{n=1}^{\infty}$ 是 $C[0, 1]$ 中的无界集,则 $\|Dx_n\|=n\not\leqslant\|D\|\|x_n\|=\|D\|$,

所以 $D:C^{(1)}[0, 1]\to C[0, 1]$是无界线性算子.

注意 若在 $C^{(1)}[0, 1]$中定义 $\|x\|'=\max\limits_{0\leqslant t\leqslant 1}|x(t)|+\max\limits_{0\leqslant t\leqslant 1}|x'(t)|$，则易证

$$D:C^{(1)}[0, 1]\to C[0, 1]$$

是有界线性算子.

定义 3 **设 X, Y 是赋范线性空间,记 $B(X, Y)=\{A|A:X\to Y$ 是有界线性算子$\}$. 如果 $A, B\in B(X, Y)$, $\alpha\in K$, $\forall x\in X$ 有**

$$(A+B)x=Ax+Bx, \ (\alpha A)x=\alpha Ax,$$

且定义 $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|$，则说 $B(X, Y)$是 X 到 Y 中的有界线性算子空间.

1. $B(X, Y)$是线性空间.

(1) $\forall A, B\in B(X, Y)$, $\forall\alpha, \beta\in K$, $\forall x\in X, y\in X$, $a, b\in K$. 易证 $(\alpha A+\beta B)(ax+by)=a(\alpha A+\beta B)x+b(\alpha A+\beta B)y$.

(2) 有界性: $\forall A, B\in B(X, Y)$,因为

$$\|(A+B)x\|=\|Ax+Bx\|\leqslant\|Ax\|+\|Bx\|$$
$$\leqslant(\|A\|+\|B\|)\|x\|, \quad \forall x\in X;$$
$$\|(\alpha A)x\|=\|\alpha Ax\|=|\alpha|\ \|Ax\|\leqslant|\alpha|\ \|A\|\ \|x\|,$$

所以 $\forall A, B\in B(X, Y)$,有 $A+B, \alpha A$ 有界.

故 $A+B, \alpha A\in B(X, Y)$,于是 $B(X, Y)$是线性空间.

2. $\|A\|$ 是 $B(X, Y)$ 上的范数.

(1) 因为 $\|A\|\geqslant 0$，且

$$\|A\| = \sup_{\|x\|=1}\|Ax\| = 0 \Leftrightarrow \|Ax\| = 0$$
$$\Leftrightarrow Ax = \theta \Leftrightarrow A = O.$$

(2) $\|\alpha A\| = \sup_{\|x\|=1}\|\alpha Ax\| = |\alpha| \sup_{\|x\|=1}\|Ax\| = |\alpha|\ \|A\|$.

(3) $\|A+B\| = \sup_{\|x\|=1}\|(A+B)x\| \leqslant \sup_{\|x\|=1}(\|Ax\| + \|Bx\|)$
$\leqslant \sup_{\|x\|=1}\|Ax\| + \sup_{\|x\|=1}\|Bx\| = \|A\| + \|B\|$.

故 $B(X, Y)$ 是赋范线性空间.

特别地，当 $Y = X$ 时，记 $B(X, X) = B(X)$.

定理 5 **若 Y 是 Banach 空间，则 $B(X, Y)$ 是 Banach 空间.**

证 设 $\{A_n\}$ 为 $B(X, Y)$ 中的 Cauchy 列，

(1) $\forall \varepsilon > 0$，$\exists N$，当 $m > n > N$，有 $\|A_m - A_n\| < \varepsilon$，于是固定 $x \in X$，有

$$\|A_m x - A_n x\| \leqslant \|A_m - A_n\|\ \|x\| \leqslant \varepsilon \|x\|. \tag{4}$$

于是 $\{A_n x\}$ 是 Y 中的 Cauchy 列，因为 Y 是 Banach 空间，所以 $\exists Ax \in Y$，使得 $\lim_{n\to\infty} A_n x = Ax$. 下面证 A 是线性算子.

$\forall x_1, x_2 \in X$，$\alpha, \beta \in K$. 因为

$$A(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} A_n(\alpha x_1 + \beta x_2) = \lim_{n\to\infty}(\alpha A_n x_1 + \beta A_n x_2)$$
$$= \alpha A x_1 + \beta A x_2.$$

故 $A: X \to Y$ 是线性算子.

(2) 证明：$A_n - A \in B(X, Y)$.

在式(4)中，令 $m\to\infty$，由范数连续性，有

$$\|A_n x - Ax\| \leqslant \varepsilon \|x\|, \qquad \forall x \in X,\ n > N.$$

即 $\dfrac{\|(A_n - A)x\|}{\|x\|} \leqslant \varepsilon$. 于是 $\|A_n - A\| \leqslant \varepsilon$，$\forall n > N$. 所以 $A_n - A$ 有界，且 $A_n - A$ 线性，故 $A_n - A \in B(X, Y)$.

(3) 因为 $A = (A - A_n) + A_n$，

$\|Ax\| \leqslant \|A - A_n\|\ \|x\| + \|A_n\|\ \|x\| \leqslant (M_1 + M_2)\|x\|$，

其中 $\|A - A_n\| \leqslant M_1$，$\|A_n\| \leqslant M_2$，所以 $A \in B(X, Y)$，

因为 $\forall n > N$，$\|A_n - A\| \leqslant \varepsilon$，所以 $\lim_{n\to\infty} A_n = A$.

三、共鸣定理

定义 4 **设 A_n，$A \in B(X, Y)$，$n = 1, 2, \cdots$**

(1) 若$\lim\limits_{n\to\infty}\|A_n-A\|=0$，则说$\{A_n\}$按范数收敛于$A$，或说$\{A_n\}$一致收敛于$A$.

(2) 若对任一固定的$x\in X$，有$\lim\limits_{n\to\infty}\|A_nx-Ax\|=0$，则说$\{A_n\}$强收敛于$A$，记为$(S)\lim\limits_{n\to\infty}A_n=A$ 或 $A_n\xrightarrow{S}A$.

定理 6 $A_n,A\in B(X,Y)$，$n=1,2,\cdots$，$\lim\limits_{n\to\infty}\|A_n-A\|=0\Rightarrow A_n\xrightarrow{S}A$. 即：一致收敛⇒强收敛.

证 因为$\|A_n-A\|\to 0$，于是对$\forall x\in X$，有

$$\|A_nx-Ax\|\leqslant\|A_n-A\|\,\|x\|\to 0,$$

所以 $A_n\xrightarrow{S}A$.

但强收敛不一定一致收敛.

例 4 A_n：$l^2\to l^2$，$\forall x=(\xi_1,\xi_2,\cdots,\xi_n\cdots)\in l^2$，定义$A_nx=(\xi_{n+1},\xi_{n+2},\cdots)$，则

(1) $A_n\in B(l^2)$，$\quad\|A_nx\|=\|(\xi_{n+1},\xi_{n+2},\cdots)\|\leqslant\sum\limits_{i=1}^{\infty}|\xi_i|^2=\|x\|$；

(2) $A_n\xrightarrow{S}O$；

事实上，$\forall x\in l^2$ 有 $\sum\limits_{i=1}^{\infty}|\xi_i|^2<+\infty$，从而$\lim\limits_{i\to\infty}|\xi_i|^2=0$，于是$\lim\limits_{i\to\infty}\xi_i=0$. $\|x\|=\left(\sum\limits_{i=1}^{\infty}|\xi_i|^2\right)^{\frac{1}{2}}$；有 $\|A_nx-\theta\|=\|A_nx\|=\sum\limits_{i=n+1}^{\infty}|\xi_i|^2\to 0$.

(3) 但是，$\|A_n\|=\sup\limits_{\|x\|=1}\|A_nx\|=1$，$\|O\|=0$. 于是

$$\|A_n-O\|=\|A_n\|=1\not\to 0.$$

注意 若$\|A_n-A\|\to 0$ $(A_n,A\in B(X,Y))$，则$\{A_n\}$是$B(X,Y)$中的有界集，即$\exists M>0$，使得$\sup\limits_{n}\|A_n\|\leqslant M$，说$\{A_n\}$是一致有界的线性算子序列.

事实上，

$$\|A_n\|\leqslant\|A_n-A\|+\|A\|<\varepsilon+\|A\|\leqslant 1+\|A\|=M.$$

在许多数学问题中，人们经常遇到的不只是单个的线性有界算子，而是一族线性有界算子，并需要讨论这一族线性有界算子在什么条件下一致有界，共鸣定理回答了这个问题.

定理 7 （共鸣定理）设X是Banach空间，Y是赋范线性空间，$\{A_\alpha\mid\alpha\in\Lambda\}\subset B(X,Y)$，则$\{\|A_\alpha\|\mid\alpha\in\Lambda\}$有界的充分必要条件是$\forall x\in X$，$\{\|A_\alpha x\|\mid\alpha\in\Lambda\}$有界.

证 必要性：由 $\{\|A_\alpha\| \mid \alpha\in\Lambda\}$ 有界，则存在$M>0$ 使 $\forall\alpha\in\Lambda$，有 $\|A_\alpha\|\leqslant M$.

$\forall x\in X$ (x 固定)，有

$$\|A_\alpha x\|\leqslant\|A_\alpha\|\,\|x\|\leqslant M\|x\|,\qquad\forall\alpha\in\Lambda,$$

即
$$\{\|A_\alpha x\|\mid\alpha\in\Lambda\}\ \text{有界}.$$

充分性：用反证法，设 $\sup\limits_{\alpha\in\Lambda}\|A_\alpha\|=+\infty$，则可证在任何闭球上，$\{\|A_\alpha x\|\mid\alpha\in\Lambda\}$ 都无界.

事实上，设 $\{\|A_\alpha x\|\mid\alpha\in\Lambda\}$ 在闭球 $\overline{B}(x_0,\gamma)$ 上有界，即 $\|A_\alpha x\|\leqslant M=$ 常数，$\forall x\in\overline{B}(x_0,\gamma)$，$\alpha\in\Lambda$ 都成立. $\forall y\in X$，$y\neq\theta$，令 $x=\dfrac{\gamma}{\|y\|}y+x_0$，则$x\in\overline{B}(x_0,\gamma)$，从而 $\|A_\alpha x\|\leqslant M$，即

$$\left\|\frac{\gamma}{\|y\|}A_\alpha y+A_\alpha x_0\right\|\leqslant M.\tag{4}$$

由假设 $\sup\limits_{\alpha\in\Lambda}\|A_\alpha x_0\|=M_0<+\infty$，从而由式(4)得

$$\|A_\alpha y\|\leqslant\frac{M+M_0}{\gamma}\|y\|,$$

由 $y\in X$ 的任意性，得 $\|A_\alpha\|\leqslant\dfrac{M+M_0}{\gamma}<+\infty$，这与$\sup\limits_{\alpha\in\Lambda}\|A_\alpha\|=+\infty$ 相矛盾.

在 X 中任取开球 $B_0=B_0(x_0,\gamma_0)$，由于 $\{\|A_\alpha x\|\mid\alpha\in\Lambda\}$ 在 $\overline{B}_0$ 上无界，则存在 A_1 及 $x_1\in B_0$，使 $\|A_1x_1\|>1$，由 A_1 的连续性，则有开球 $B_1=B_1(x_1,\gamma_1)\subset B_0$，使 $\|A_1x\|>1$，$\forall x\in B_1$，这里可取$\gamma_1<\dfrac{1}{2}\gamma_0$.

再由 $\{\|A_\alpha x\|\mid\alpha\in\Lambda\}$ 在 B_1 上无界，则存在 A_2 及 $x_2\in B_1$，使 $\|A_2x_2\|>2$.

由 A_2 的连续性，则有开球 $B_2=B_2(x_2,\gamma_2)\subset B_1$ 使 $\|A_2x\|>2$，$\forall x\in B_2$，且 $\gamma_2<\dfrac{1}{2}\gamma_1$.

继续以上作法，可得有界算子列$\{A_n\}$及一串闭球 $\{\overline{B}_n\}$，$\overline{B}_n$ 满足：$\overline{B}_1\supset\overline{B}_2\supset\cdots\supset\overline{B}_n\supset\cdots$，$\gamma_n\to0$，且有 $\|A_nx\|\geqslant n$，$\forall x\in\overline{B}_n$，$n=1,2,3\cdots$. 由于 X 是 Banach 空间，故存在惟一点 $x^*\in\bigcap\limits_{n=1}^{\infty}\overline{B}_n$，从而 $\|A_nx^*\|\geqslant n$，$n=1,2,3,\cdots$ 这与对每一个固定的 $x\in X$，$\sup\limits_{\alpha\in\Lambda}\|A_\alpha x\|$ 有界相矛盾.

注意 共鸣定理告诉我们，为了证明线性有界算子族 $\{A_\alpha\mid\alpha\in\Lambda\}$ 一致有

界，只需证明对 Banach 空间 X 中的每个 x，$\{A_\alpha x \mid \alpha \in \Lambda\}$ 在 Y 中有界即可. 因此，共鸣定理也称为一致有界定理.

推论 1 若 X 是 Banach 空间，A_n，$A \in B(X, Y)$，则当 $A_n \xrightarrow{S} A$，必有

$$\sup_n \| A_n \| < +\infty.$$

四、下有界线性算子与等价范数

定义 5 设 X，Y 是赋范线性空间，$A: X \to Y$ 是线性算子，若 $\exists C > 0$，使得 $\| Ax \| \geqslant C \| x \|$，$\forall x \in X$，则称 A 是下有界线性算子.

例 5 设 X 是赋范线性空间，$A \in B(X)$，$\exists \lambda \in K$，满足 $|\lambda| > \| A \|$，则 $\lambda I - A$ 是下有界线性算子.

证 $\forall x \in X$，有

$$\begin{aligned} \| (\lambda I - A)x \| &= \| \lambda x - Ax \| \geqslant \| \lambda x \| - \| Ax \| \\ &= |\lambda| \, \| x \| - \| Ax \| = (|\lambda| - \| A \|) \| x \|. \end{aligned}$$

例 6 设 $C_0[0, 1] = \{x \in C^{(1)}[0, 1] \mid x(0) = x(1) = 0\}$，$\| x \| = \max\limits_{0 \leqslant t \leqslant 1} | x(t) |$，则 $D = \dfrac{\mathrm{d}}{\mathrm{d}t}: C_0[0, 1] \to C[0, 1]$.

显然 D 是无界算子，但 D 是下有界的.

取 $x(t) = \sin n\pi t$，$\{\sin n\pi t\} \subset C_0[0, 1]$ 有界. 但 $D(\sin n\pi t) = \{n\pi \cos n\pi t\}$ 是$C[0, 1]$ 中的无界集.

另一方面，$\forall x \in C_0[0, 1]$，

$$\begin{aligned} | x(t) | &= \left| \int_0^1 \frac{\mathrm{d}x(t)}{\mathrm{d}t} \mathrm{d}t \right| \\ &\leqslant \max_{0 \leqslant t \leqslant 1} \left| \frac{\mathrm{d}x(t)}{\mathrm{d}t} \right| = \| Dx \|, \ t \in [0, 1]. \end{aligned}$$

故 $\| Dx \| \geqslant \| x \| = \max\limits_{0 \leqslant t \leqslant 1} | x(t) |$，即 D 下有界.

定理 8 （下有界线性算子特征）$A: X \to Y$ 为下有界线性算子，则 $A^{-1} \in B(R(A), X)$. 反之，$A: X \to Y$ 为线性算子，且 $A^{-1} \in B(R(A), X)$，则 A 是下有界的.

定义 6 X，Y 是 K 上赋范线性空间，若 $\exists H \in B(X, Y)$，且 $H^{-1} \in B(Y, X)$，则说 H 是 X 到 Y 上的拓扑同构映射，X 与 Y 拓扑同构.

定理 9 设 $H: (X, \| \cdot \|) \to (Y, \| \cdot \|)$ 为线性算子，则 H 为拓扑同构映射的充分必要条件是 H 是满射，且 $\exists \mu_1, \mu_2 > 0$，使得 $\forall x \in X$，有 $\mu_1 \| x \| \leqslant \| Hx \| \leqslant \mu_2 \| x \|$.

定义 7 设 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 X 上的两个范数,若恒等映射 $I:(X,\ \|\cdot\|_1)\to(X,\ \|\cdot\|_2)$ 是拓扑同构映射,则说 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 X 上的等价范数.

定理 10 $\|\cdot\|_1$, $\|\cdot\|_2$ 是 X 上的范数,则 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价的充分必要条件是 $\exists \mu_1, \mu_2>0$, $\forall x\in X$, 有

$$\mu_1\|x\|_1\leqslant\|x\|_2\leqslant\mu_2\|x\|_1.$$

例 7 $\mathbf{E}^n$ 上的范数 $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ 等价,其中

$$\|x\|_1=\sum_{i=1}^n|\xi_i|,\qquad \|x\|_2=\left(\sum_{i=1}^n|\xi_i|^2\right)^{\frac{1}{2}},$$

$$\|x\|_\infty=\max_{1\leqslant i\leqslant n}|\xi_i|.$$

证 $\forall x\in\mathbf{E}^n$,

$$\|x\|_\infty=\max_{1\leqslant i\leqslant n}|\xi_i|\leqslant\sum_{i=1}^n|\xi_i|=\|x\|_1\leqslant n\|x\|_\infty,$$

$$\|x\|_\infty=\max_{1\leqslant i\leqslant n}|\xi_i|\leqslant\left(\sum_{i=1}^n|\xi_i|^2\right)^{\frac{1}{2}}=\|x\|_2\leqslant\sqrt{n}\|x\|_\infty,$$

所以 $\|\cdot\|_1$, $\|\cdot\|_\infty$, $\|\cdot\|_2$ 相互等价.

定理 11 有限维赋范线性空间 X 上的线性算子都是有界的.

证 设 $\dim X=n$, X 的基为 $\{e_1, e_2, \cdots, e_n\}$.

$A:X\to Y$ 线性算子, $\forall x\in X$, 有 $x=\sum\limits_{i=1}^n\xi_ie_i$,

$$Ax=A\left(\sum_{i=1}^n\xi_ie_i\right)=\sum_{i=1}^n\xi_iAe_i\in Y,$$

于是

$$\|Ax\|\leqslant\sum_{i=1}^n\|\xi_iAe_i\|=\sum_{i=1}^n|\xi_i|\ \|Ae_i\|$$

$$\leqslant(|\xi_1|+|\xi_2|+\cdots+|\xi_n|)\max_{1\leqslant i\leqslant n}\|Ae_i\|.$$

设 f_i: $X\to K$ 线性泛函,满足:

$$f_i(e_j)=\delta_{ij}=\begin{cases}0, & \text{当 } i\neq j \text{ 时};\\ 1, & \text{当 } i=j \text{ 时},\end{cases}\quad i, j=1, 2, \cdots, n,$$

则

$$\forall x=\sum_{j=1}^n\xi_je_j\in X \text{ 有 } f_i(x)=\sum_{j=1}^n\xi_j f_i(e_j)=\xi_i,$$

易验证 f_i 连续,由本章第二节定理 3 知 f_i 有界,$i=1, 2, \cdots, n$. 于是 $\forall x\in X$,

$$|\xi_i|=|f_i(x)|=\|f_i(x)\|\leqslant\|f_i\|\ \|x\|,$$

所以 $\| Ax \| \leqslant (\| f_1 \| + \| f_2 \| + \cdots + \| f_n \|) \| x \| \max\limits_{1\leqslant j\leqslant n} \| Ae_j \|$

$$= M \| x \| .$$

于是 $A \in B(X, Y)$.

定理 12 **有限维空间 X 上的范数都等价. n 维实(复)赋范线性空间都与 $\mathbf{E}^n(\mathbf{C}^n)$ 拓扑同构.**

推论 2 **有限维赋范线性空间 X 必为 Banach 空间,且与 $\mathbf{E}^n$ 拓扑同构.**

推论 3 **赋范线性空间的有限维子空间必为闭子空间.**

Riesz 引理:

设 E 是赋范线性空间 X 的真闭子空间则 $\forall \varepsilon > 0, 0 < \varepsilon < 1$ 必存在 $x_0 \in X$, 使得 $\| x_0 \| = 1$, 且 $d(x_0, E) > \varepsilon$.

定理 13 **赋范线性空间 X 是有限维的充分必要条件是 X 中紧集等价于有界闭集.**

第三节　有界线性泛函与共轭空间

一、共轭空间

定义 1 **设 $X^* = B(X, K)$ 是 X 上全体有界线性泛函所组成的赋范线性空间.**

$$f \in X^*, \ \| f \| := \sup_{x\neq\theta} \frac{| f(x) |}{\| x \|} = \sup_{\| x \| =1} | f(x) | = \sup_{\| x \| \leqslant 1} | f(x) | ,$$

则称 X^* 为赋范线性空间 X 的共轭空间.

例 1 $f: C[a, b] \to \mathbf{R}$, $f(x) = \int_a^b x(t)\mathrm{d}t$, $x \in C[a, b]$, 则 $f \in (C[a, b])^*$, 且 $\| f \| = b - a$.

证 (1) 显然 f 是线性泛函.

(2) $\forall x \in C[a, b]$,

$$| f(x) | = \left| \int_a^b x(t)\mathrm{d}t \right| \leqslant \int_a^b | x(t) | \mathrm{d}t$$

$$\leqslant (b - a) \max_{a\leqslant t\leqslant b} | x(t) | = (b - a) \| x \| .$$

故 f 有界,从而 $f \in (C[a, b])^*$ 且 $\dfrac{| f(x) |}{\| x \|} \leqslant b - a$, 则 $\| f \| \leqslant b - a$.

另一方面,取 $x_0(t) \equiv 1 \Rightarrow \| x_0 \| = 1$, 且

$$0 < b - a = \int_a^b x_0(t)\mathrm{d}t = | f(x_0) | \leqslant \| f \| \| x_0 \| = \| f \| ,$$

所以 $\|f\| = b-a$.

定义 2 设 X, Y 是两个赋范线性空间，$A: X \to Y$ 为线性算子，若

(1) A 是满射；

(2) A 是保范的，即 $\|AX\| = \|x\|$，$\forall x \in X$，

则说 A 是 X 到 Y 上的保范同构映射.

若 X 与 Y 之间存在保范同构映射，则说 X 与 Y 是保范同构的.

二、Hahn-Banach 定理

X 上是否一定存在非零的有界线性泛函？这是泛函分析中的最基本问题.

定理 1 （Hahn-Banach 延拓定理）设 M 是赋范线性空间 X 的子空间，$f \in M^*$，则 f 可以保范地延拓到整个空间 X 上，即 $\exists F \in X^*$，

(1) $F(x) = f(x)$，$\forall x \in M$；

(2) $\|F\| = \|f\|$.

定理 2 X 是赋范线性空间，M 是 X 的子空间，$x_0 \in X$，且 $d(x_0, M) = d > 0$，则 $\exists f \in X^*$，使得

(1) $f(x) = 0$，$\forall x \in M$；

(2) $f(x_0) = d$；

(3) $\|f\| = 1$.

证 (1) 令 $M_1 = \{x + \alpha x_0 \mid x \in M, \alpha \in K\}$，则 M_1 是 X 的子空间，并且 $x_0 \notin M$，$M_1 \supset M$（因为 $x_0 \in M$ 时 $d(x_0, M) = 0$）.

再令 $f_0: M_1 \to K$，$f_0(x + \alpha x_0) = \alpha d$，$\forall x + \alpha x_0 \in M_1$，则 f_0 是线性泛函，且 $f_0(x_0) = d$，$f_0(x) = 0$，$\forall x \in M$.

由 $x \in M$，M 是 X 的线性子空间可知 $-\frac{1}{\alpha}x \in M$，于是$\left\|x_0 - \left(-\frac{x}{\alpha}\right)\right\| \geqslant d(x_0, M) = d$. 所以

$$|f_0(x + \alpha x_0)| = |\alpha| d \leqslant |\alpha| \left\|x_0 - \left(-\frac{x}{\alpha}\right)\right\|$$

$$= |\alpha| \left\|\frac{\alpha x_0 + x}{\alpha}\right\| = \|\alpha x_0 + x\|,$$

即
$$\frac{|f_0(x + \alpha x_0)|}{\|\alpha x_0 + x\|} \leqslant 1,$$

因此 $\|f_0\| \leqslant 1$，即 f_0 是 M_1 上的有界线性泛函，$f_0 \in M_1^*$.

另一方面，由 $d = d(x_0, M)$ 定义，$\forall n$，$\exists x_n \in M$，使得

$$\|x_n - x_0\| < d + \frac{1}{n}.$$

$$\| f_0 \| \| x_n - x_0 \| \geqslant \| f_0(x_n - x_0) \| = |-d| = d, \ \forall n.$$

于是
$$\| f_0 \| \geqslant \frac{d}{\| x_n - x_0 \|} \geqslant \frac{d}{d + \frac{1}{n}} \to 1.$$

所以 $\| f_0 \| \geqslant 1$，从而 $\| f_0 \| = 1$. 因 f_0 是 M 上的有界线性泛函，再用 Hahn-Banach 定理，f_0 可保范延拓成 X 上的有界线性泛函 $f \in X^*$，

(1) $f(x) = f_0(x) = 0, \ x \in M_1$；

(2) $f_0(x_0) = f(x_0) = d$；

(3) $\| f \| = \| f_0 \| = 1$.

三、弱收敛

定义 3 设 $x_n, x_0 \in (X, \| \cdot \|)$，$\forall f \in X^*$，有 $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$.

则说 $\{x_n\}$ **弱收敛**于 x_0，记为 $\lim\limits_{n\to\infty} x_n \overset{w}{=\!=} x_0$ 或 $x_n \xrightarrow{w} x_0$.

定义 4 设 $f_n, f_0 \in X^*$，

(1) 若 $\lim\limits_{n\to\infty} \| f_n - f_0 \| = 0$，则说 $\{f_n\}$ 强收敛于 f_0，记为

$$\lim_{n\to\infty} f_n \overset{S}{=\!=} f_0 \text{ 或 } f_n \xrightarrow{S} f_0$$

(2) 若 $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$，　$\forall x \in X$，则说 $\{f_n\}$ 弱*收敛于 f_0，记为 $\lim\limits_{n\to\infty} f_n \overset{w^*}{=\!=} f_0$ 或 $f_n \xrightarrow{w^*} f_0$.

定理 3 （弱收敛的性质）设 $x_n, x_0 \in (X, \| \cdot \|)$，

(1) 设 $\{x_n\}$ 弱收敛，则其弱极限 x_0 是惟一的；

(2) 若 $\{x_n\}$ 弱收敛，则 $\{x_n\}$ 在 X 中有界，即 $\{\| x_n \|\}$ 为有界数列；

(3) $x_n \xrightarrow{w} x_0$ 的充分必要条件是 $\{\| x_n \|\}$ 为有界数列，且对 X^* 的稠密子集 M 上每个元素 f 都有 $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$.

第四节　闭图像定理与有界逆算子定理

设 X, Y 是 K 上的赋范线性空间，在 $X \times Y$ 上定义如下线性运算：

$$\forall x, x_1, x_2 \in X, \qquad y, y_1, y_2 \in Y,$$

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$

$$\alpha(x, y) = (\alpha x, \alpha y), \quad \forall \alpha \in K.$$

$$\|(x, y)\| = (\|x\|^p + \|y\|^p)^{\frac{1}{p}}, \quad 1 \leqslant p < +\infty.$$

则易证,$(X \times Y, \|\cdot\|)$是赋范线性空间,简称 X 与 Y 的积空间.

定义 1 设 X, Y 为赋范线性空间,$A: D(A) \subseteq X \to Y$ 是线性算子,且 $D(A)$ 是 X 的子空间,若 A 的图像 $\mathrm{Graph}(A) = \{(x, Ax) \mid x \in D(A)\}$ 是 $X \times Y$ 中的闭集,则称 A 为闭线性算子,简称闭算子.

注意 (1) 线性算子不一定是有界的.

(2) 有些无界算子是闭算子.

定理 1 线性算子 $A: D(A) \subseteq X \to Y$ 是闭的充分必要条件是 $\forall \{x_n\} \subseteq D(A)$,若 $x_n \to x_0$,且 $Ax_n \to y_0$,则必有 $x_0 \in D(A)$ 且 $Ax_0 = y_0$.

证 必要性:A 是闭算子,则 $\mathrm{Graph}(A)$是 $X \times Y$ 中的闭集.

由 $x_n \to x_0$, $Ax_n \to y_0$ 知$(x_n, Ax_n) \to (x_0, y_0)$.

充分性:$\forall (x_0, y_0) \in \overline{\mathrm{Graph}(A)}$,下证$(x_0, y_0) \in \mathrm{Graph}(A)$.

$\exists x_n \in D(A)$,使得$(x_n, Ax_n) \to (x_0, y_0)$,即 $x_n \to x_0$, $Ax_n \to y_0$. 由条件知$x_0 \in D(A)$, $Ax_0 = y_0$,从而$(x_0, y_0) = (x_0, Ax_0) \in \mathrm{Graph}(A)$.

所以 Graph(A)是闭集,A 是闭算子.

例 1 微分算子 $D = \frac{\mathrm{d}}{\mathrm{d}t}: C^{(1)}[0, 1](\subset C[0, 1]) \to C[0, 1]$. 由本章第二节例 6 知,$D$ 是无界线性算子,但 D 是闭算子.

证 设 $\{x_n\} \subset C^{(1)}[0, 1]$,使得 $x_n \to x_0$, $Dx_n \to y_0$. 这里显然是$[0, 1]$上函数序列的一致收敛.

$\forall t \in [0, 1]$,

$$\begin{aligned}
\int_0^t y_0(\tau)\mathrm{d}\tau &= \int_0^t \lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}\tau} x_n(\tau)\mathrm{d}\tau \\
&= \lim_{n\to\infty} \int_0^t \frac{\mathrm{d}x_n(\tau)}{\mathrm{d}\tau}\mathrm{d}\tau \\
&= \lim_{n\to\infty} [x_n(t) - x_n(0)] \\
&= x_0(t) - x_0(0).
\end{aligned}$$

于是 $x_0(t) = x_0(0) + \int_0^t y_0(\tau)\mathrm{d}\tau$. 由 $x'_0(t) = y_0(t)$ 知 $x_0(t) \in C^{(1)}[0, 1]$,并且 $Dx'_0 = y_0$. 由定理 1 知 D 是闭算子.

定理 2 (闭图像定理)若 X, Y 是 Banach 空间,$A: X \to Y$ 为线性算子. 则 A 有界的充分必要条件是 A 是闭算子.

注意 判定线性算子连续性(有界性)转为判定闭性更容易. 因为线性算子 $A: X \to Y$,对下面三条

(1) $x_n \to x_0$； (2) $Ax_n \to y_0$； (3) $y_0 = Ax_0$

若判定 A 在 x_0 连续，则需由(1)推导出(2)和(3). 若判定 A 闭，则由(1)和(2)推导出(3).

定理 3 (Banach 有界逆算子定理)**设 X, Y 是 Banach 空间，$A:X\to Y$ 是有界线性算子，且为双射，则 $A^{-1} \in B(Y, X)$.**

证 因为 $A:X\to Y$ 是双射的有界线性算子. 则 $A^{-1}:Y\to X$ 存在且是线性算子.

下证 A^{-1}是闭算子：

事实上，y_n, $y_0 \in Y$，若 $y_n \to y_0$ 且 $A^{-1}y_n \to x_0$，由 A 连续知 $A(A^{-1}y_n) \to Ax_0$，即 $y_n \to Ax_0$，故 $y_0 = Ax_0$(度量空间中极限惟一)，即 $x_0 = A^{-1}y_0 \in X$. 由定理 1 知 A^{-1} 是闭算子. 由闭图像定理知 A^{-1} 有界，所以 $A^{-1} \in B(Y, X)$.

推论 1 **X, Y 是 Banach 空间，$A:X\to Y$ 是单射的有界线性算子，则 A 的值域$R(A)$是 Y 的闭子空间的充分必要条件是A下有界.**

证 必要性：设 $R(A)$是 Banach 空间 Y 的闭子空间，故$R(A)$是Banach空间，由本节定理 3 知$A^{-1}\in B(R(A), X)$.

再由本章第二节定理 8 知 A 是下有界的.

充分性：设 $A:X\to Y$ 是下有界的有界线性算子，故 $\exists C>0$，使得 $C\|x\| \leqslant \|Ax\| \leqslant \|A\| \|x\|$，$\forall x \in X$. 由本章第二节定理 9 知 X 与$R(A)$是拓扑同构，从而X与$R(A)$有相同的完备性.

由 X 完备可知$R(A)$完备，从而$R(A)$是 Y 的闭子空间.

第五节 紧 算 子

定义 **设 X, Y 是赋范线性空间，$A:X\to Y$是线性算子，若 $\forall M\subseteq X$ 有界，有$A(M) \subseteq Y$ 中列紧，则称 A 是紧算子(或全连续算子).**

注意 (1) 线性算子 $A:X\to Y$ 是紧算子的充分必要条件是 $A(V[\theta, 1])$是 Y 中列紧集，其中 $V[\theta, 1] = \{x \mid \|x\| \leqslant 1\}$ 为 X 中单位闭球. A 是紧算子必有 A 是连续的，反之未必.

(2) $A\in B(X, Y)$是紧算子的充分必要条件是X中任一有界点列$\{x_n\}$，$\{Ax_n\}$总存在收敛子列.

(3) 若 $A \in B(X, Y)$，且 $\dim R(A) <+\infty$，则说 A 是有限秩算子，有限秩算子 A 必是紧算子，其中 $\dim R(A)$表示 $R(A)$的维数.

事实上，若 A 是有限秩算子，则$R(A)$是 Y 的有限维子空间. 对 X 中任一有界集M，$A(M)$是$R(A)$中的有界集. 而 $R(A)$有限维，所以$A(M)$是列紧，故 A 是紧算子.

结论：X, Y 是赋范线性空间，$\dim X <+\infty$，线性算子 $A:X\to Y$ 必是紧算子.

例 恒等算子 $I:l^2\to l^2$ 不是紧算子，而是线性算子.

令 $M=\{e_n\}_{n=1}^{\infty}$， $e_n=(0,0,\cdots,1,0,\cdots)$. 则

(1) M 是 l^2 中的有界集；

(2) 若 $A\in B(X,Y)$，$B\in B(Y,Z)$，且 A，B 中有一个是紧算子，则BA是紧算子；

(3) 当 Y 是 Banach 空间时，$C(X,Y)$是 Banach 空间.

第六节 压缩映射定理

一、压缩映射定理

定义 1 (1) 设 $F:X\to X$，若 $\exists x^*\in X$，有 $F(x^*)=x^*$，则说 x^* 是 F 的一个不动点.

(2) (X,d)是度量空间，$\exists 0\leqslant k<1$，使得 $\forall x,y\in X$，有

$$d(F(x),F(y))\leqslant kd(x,y),$$

则说 $F:X\to X$ 是压缩映射.

注意 (1) 压缩映射是连续的.

(2) 一般映射未必有不动点，若有也未必惟一.

定理 1 (Banach 压缩映射原理)设(X,d)是完备度量空间，$F:X\to X$ 是压缩映射，则 $\exists | x^*\in X$，使得$F(x^*)=x^*$.

证 (1) 存在性：$\forall x_0\in X$，令 $x_1=F(x_0)$，$x_2=F(x_1)$，$\cdots$，$x_n=F(x_{n-1})$，下证$\{x_n\}$是 X 中的 Cauchy 序列.

事实上，由 $d(x_{r+1},x_r)=d(F(x_r),F(x_{r-1}))$

$$\leqslant k(x_r,x_{r-1})\leqslant\cdots\leqslant k^r d(x_1,x_0),$$

故当 $m>n$ 时，

$$\begin{aligned}
d(x_m,x_n)&=d(F(x_{m-1}),F(x_{n-1}))\leqslant kd(x_{m-1},x_{n-1})\\
&\leqslant k^2d(x_{m-2},x_{n-2})\leqslant\cdots\leqslant k^nd(x_{m-n},x_0)\\
&\leqslant k^n[d(x_{m-n},x_{m-n-1})+d(x_{m-n-1},x_{m-n-2})+\cdots+d(x_1,x_0)]\\
&\leqslant k^n[k^{m-n-1}+k^{m-n-2}+\cdots+k+1]d(x_1,x_0)\\
&\leqslant k^n\cdot\frac{1-k^{m-n}}{1-k}d(x_1,x_0)\leqslant\frac{k^n}{1-k}d(x_1,x_0),
\end{aligned}$$

由 $0\leqslant k<1$ 知$\lim\limits_{n\to\infty}\frac{k^n}{1-k}d(x_1,x_0)=0$.

故$\{x_n\}$是 X 中的 Cauchy 序列. 而由 X 完备知$\exists x^* \in X$,使得 $\lim\limits_{n\to\infty} x_n = x^*$.
又由 F 连续知 $\lim\limits_{n\to\infty} x_{n+1} = \lim\limits_{n\to\infty} F(x_n) = F(x^*) = x^*$.

(2) 惟一性:若 $\exists x^*, y^* \in X$,有$F(x^*) = x^*$, $F(y^*) = y^*$. 则

$$d(x^*, y^*) = d(F(x^*), F(y^*)) \leqslant kd(x^*, y^*).$$

而 $0 \leqslant k < 1$ 知 $d(x^*, y^*) = 0$,故 $x^* = y^*$.

注意 (1) 在完备度量空间中,可用迭代法求压缩映射不动点. $\forall x_0 \in X$,令 $x_n = F(x_{n-1})(n = 1, 2, \cdots)$. 其误差 $d(x_n, x^*) \leqslant \dfrac{k^n}{1-k} d(x_1, x_0)$.

(2) 定理 1 中 X 的完备性与 F 的压缩性是不可缺少的条件.

定理 2 (X, d)**是完备度量空间,且** $F^n = F \cdot F \cdots F$: $X \to X$ **是压缩映射,则** $F: X \to X$ **有惟一不动点** x^*.

证 (1) 由定理 1, $\exists | x^* \in X$, 使得 $F^n(x^*) = x^*$,由 $F^n(F(x^*)) = F^{n+1}(x^*) = F(F^n(x^*)) = F(x^*)$ 知$F(x^*)$也是 F^n 的惟一不动点.

故 $x^* = F(x^*)$, 即 x^* 是 F 的不动点.

(2) 若 $y^* = F(y^*)$ 则 $F^n(y^*) = F^{n-1}(F(y^*)) = F^{n-1}(y^*) = \cdots = y^*$,由 F^n 的不动点的惟一性知$x^* = y^*$.

二、应用举例

例 1 (方程的实根的存在性)设 $F: \mathbf{R} \to \mathbf{R}$,且 $F(x)$在 $\mathbf{R}$ 上可导,并满足 $| F'(x) | \leqslant \alpha < 1$,则方程$F(x) = x$ 存在实根.

证 由中值定理知 $| F(x) - F(y) | = | F'(\xi)(x-y) | \leqslant \alpha | x - y |$,其中 ξ 在 x 与 y 之间,即

$$d(F(x), F(y)) \leqslant \alpha d(x, y), \qquad 0 \leqslant \alpha < 1.$$

故 F 是压缩映射. 于是由 $\mathbf{R}$ 完备知方程 $F(x) = x$ 有惟一解 x^*.

$\forall x_0 \in R$,令 $x_n = F(x_{n-1})$,则得近似解 x_n 的误差估计:

$$d(x_n, x^*) \leqslant \frac{\alpha^n}{1-\alpha} d(F(x_0), x_0).$$

即

$$| x_n - x^* | \leqslant \frac{\alpha^n}{1-\alpha} | F(x_0) - x_0 |.$$

例 2 F-积分方程 $$x(t) = f(t) + \lambda \int_a^b k(t, \tau) x(\tau) d\tau, \tag{1}$$

其中 $k(t, \tau)$ 是矩形域$A = [a, b] \times [a, b]$ 上连续函数; $f(t) \in C[a, b]$ 且 $\forall t \in [a, b]$ 有

$$\int_a^b |k(t,\tau)|\,\mathrm{d}\tau \leqslant M < +\infty;$$

而 $x(t)$是未知函数,则方程(1)当 $|\lambda| < \frac{1}{M}$ 时,有惟一解$\tilde{x} \in C[a, b]$.

证 令 $Tx = f(t) + \lambda\int_a^b k(t,\tau)x(\tau)\mathrm{d}\tau$,则 $T: C[a, b] \to C[a, b]$. $\forall x, y \in C[a, b]$ 有

$$\begin{aligned}
\mathrm{d}(Tx, Ty) &= \max_{t\in[a,b]} |Tx(t) - Ty(t)| \\
&= \max_{t\in[a,b]} \left| f(t) + \lambda\int_a^b k(t,\tau)x(\tau)\mathrm{d}\tau - f(t) - \lambda\int_a^b k(t,\tau)y(\tau)\mathrm{d}\tau \right| \\
&\leqslant |\lambda| \max_{t\in[a,b]}\int_a^b |k(t,\tau)||x(\tau) - y(\tau)|\,\mathrm{d}\tau \\
&\leqslant |\lambda| \cdot M \cdot \max_{\tau\in[a,b]} |x(\tau) - y(\tau)| \\
&= |\lambda| Md(x, y).
\end{aligned}$$

由于 $|\lambda| M < 1$, $C[a, b]$ 是完备的,因此由 Banach 压缩映射定理知,积分方程(1)有惟一的连续解 $\tilde{x}(t)$.

习 题 3-1

1. 下列集合,哪些是线性空间?

(1) 满足 $\lim\limits_{n\to\infty} \xi_n = 0$ 的数列 $x = (\xi_1, \cdots, \xi_n, \cdots)$ 的全体;

(2) 满足 $\lim\limits_{n\to\infty} \xi_n = 1$ 的数列 $x = (\xi_1, \cdots, \xi_n, \cdots)$ 的全体;

(3) 满足微分方程 $\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = 0$ (b, c 为已知实数) 的一切解 $y = y(x)$ 的全体;

(4) 一切 x 的多项式的全体;

(5) 一切次数≤4 的 x 的多项式的全体.

2. 设 M 为线性空间 X 的子集,证明:

(1) span$M$ 是 X 的子空间;

(2) span$M$ 是 X 的包含 M 的最小子空间.

3. 设 Y, Z 都是线性空间 X 的子空间,证明:$Y\cap Z$是 X 的子空间. 试问:$Y\cup Z$是否也一定是 X 的子空间?

4. 设 $Af = \int_a^x f(t)\mathrm{d}t$, $f \in C[a, b]$, $Df = \frac{\mathrm{d}}{\mathrm{d}x}f(x)$, $f \in C^{(1)}[a, b]$.

证明:$A: C[a, b] \to C^{(1)}[a, b]$ 是线性算子.

$D: C^{(1)}[a, b] \to C[a, b]$ 是线性算子.

5. 设 $A:D(A)(\subset X)\to Y$ 是线性算子，证明：$D(A)$ 为 X 的子空间；$R(A)$ 为 Y 的子空间.

6. 若 $\dim X=\dim Y<+\infty$，$A:X\to Y$ 为线性算子，证明：A^{-1} 存在的充分必要条件是 $R(A)=Y$.

7. 设 $\{A_\lambda \mid \lambda\in\Lambda\}$ 为线性空间 X 中的一族凸集，证明：$\bigcap\limits_{\lambda\in\Lambda}A_\lambda$ 也是凸集.

8. 证明：$\|x\|_\infty=\max\limits_{1\leqslant i\leqslant n}|\xi_i|$ $(x=(\xi_1,\cdots,\xi_n)\in \mathbf{E}^n)$ 是 $\mathbf{E}^n$ 上的范数.

9. 证明：$\|x\|_p=\left(\int_a^b|x(t)|^p\mathrm{d}t\right)^{1/p}$ $(1\leqslant p<+\infty)$ 是 $C[a,b]$ 上的范数.

10. 设 X 为Banach空间，证明：

(1) 若 $\sum\limits_{n=1}^{\infty}\|x_n\|$ 收敛，则 $\sum\limits_{n=1}^{\infty}x_n$ 收敛；

(2) 若 $\sum\limits_{n=1}^{\infty}x_n$ 收敛，则 $\lim\limits_{n\to\infty}x_n=\theta$；

(3) 若 $\sum\limits_{n=1}^{\infty}x_n$ 收敛，则 $\left\|\sum\limits_{n=1}^{\infty}x_n\right\|\leqslant\sum\limits_{n=1}^{\infty}\|x_n\|$.

11. 若 X 为赋范线性空间，$\sum\limits_{n=1}^{\infty}\|x_n-x_{n+1}\|$ 收敛，$x_n\in X$，试问：$\{x_n\}$ 是否为Cauchy列？$\{x_n\}$ 是否一定收敛？

12. 设 X 为赋范线性空间，令

$$d(x,y)=\begin{cases}0, & \text{当 } x=y \text{ 时；}\\ \|x-y\|+1, & \text{当 } x\neq y \text{ 时.}\end{cases}$$

证明：(X,d) 为度量空间，但 d 不是由范数决定的度量.

13. 在 $C^{(1)}[0,1]$ 中定义范数 $\|x\|=\max\limits_{t\in[0,1]}|x(t)|+\max\limits_{t\in[0,1]}|x'(t)|$，证明：微分算子 $D=\dfrac{\mathrm{d}}{\mathrm{d}t}:C^{(1)}[0,1]\to C[0,1]$ 为有界线性算子.

14. 设 $A=\begin{pmatrix}2&1\\1&2\end{pmatrix}:\mathbf{E}^2\to\mathbf{E}^2$，证明：$A$ 为有界线性算子，并求 $\|A\|$.

15. 设 $A_n, A\in B(X,Y)$，则 $\|A_n-A\|\to 0$ 的充分必要条件是 $\{A_n\}$ 在 X 的任一有界集 Ω 上都一致收敛于 A.

16. 设 X 为 Banach 空间，$A\in B(X)$，$\|A\|<|\lambda|$，证明：

$$\|(\lambda I-A)^{-1}\|\leqslant\frac{1}{|\lambda|-\|A\|}.$$

17. 设 $g(x)=x(t_0)$，$x\in C[a,b]$，t_0 为 $[a,b]$ 中固定点，证明：$g\in(C[a,b])^*$，且 $\|g\|=1$.

18. 设 $f(x)=\int_a^{\frac{a+b}{2}}x(t)\mathrm{d}t$，$x\in C[a,b]$，证明：$f\in(C[a,b])^*$，并求 $\|f\|$.

19. 设 f 为赋范线性空间 X 上的线性泛函，证明：$f\in X^*$ 的充分必要条件是 $N(f)=\{x\in X;\ f(x)=0\}$ 为 X 的闭子空间.

20. 设 X 为赋范线性空间，$x_n, x\in X$，证明：当 $x_n\xrightarrow{\|\cdot\|}x$ 时必有 $x_n\xrightarrow{w}x$.

21. 设 $A:D(A)(\subset X)\to Y$ 是有界线性算子，且 $D(A)$ 为 X 的闭子空间，证明：A 是闭算子.

22. 设 $A:D(A)(\subset X)\to Y$ 是闭算子，证明：$N(A)$ 必为 X 的闭子空间.

23. 设 $(X, \|\cdot\|_1)$，$(X, \|\cdot\|_2)$ 都是 Banach 空间，且存在常数 $M>0$，使得

$$\|x\|_2 \leqslant M\|x\|_1, \ \forall x \in X.$$

证明：$\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 X 上的等价范数.

24. 设 $A:X\to Y$ 是线性算子，X, Y 为赋范空间，证明：A 为紧算子的充分必要条件是 $A(V[\theta, 1])$ 是 Y 中列紧集，其中 $V[\theta, 1]\subset X$.

25. 证明：定义在有限维赋范空间 X 到任一赋范空间 Y 中取值的线性算子必为紧算子.

第二章　Hilbert 空间

Hilbert 空间是一类具有较强几何特性的赋范线性空间，它把欧氏空间推广到无限维的情形，在数学和物理的许多领域占有重要的地位. 由于 Hilbert 空间有正交系和正交分解定理，向量可展为广义 Fourier 级数，其基本特征是在线性空间中引进了内积，因此，Hilbert 空间理论是泛函分析中最早成熟的，也是内容最丰富、应用最广泛的部分.

第一节　内积空间与 Hilbert 空间

一、内积空间的概念及特征

定义 1　X **是数域** K(**R 或 C**)**上的线性空间，若有泛函**

$$(\cdot,\ \cdot): X \times X \to K$$

满足如下公理：

1°**正定性：**　$(x,\ x) \geqslant 0$，$\forall\, x \in X$，且$(x,\ x) = 0 \Leftrightarrow x = \theta$.

2°**共轭对称：**　$(x,\ y) = \overline{(y,\ x)}$　$\forall\, x,\ y \in X$.

3°**对第一变元线性：**

$(\alpha x + \beta y,\ z) = \alpha(x,\ z) + \beta(y,\ z)$　$\forall\, x,\ y,\ z \in X,\ \alpha,\ \beta \in K$.

则说$(\cdot,\ \cdot)$**是** X **中内积.**

定义了内积的线性空间 X 称为内积空间. 当 K 是 **R**(或 **C**)时，则 X 称为实(复)内积空间.

注意　$\bar{z} \cdot z = |z|^2$，$\overline{z_1 + z_2} = \bar{z_1} + \bar{z_2}$，$\overline{z_1 \cdot z_2} = \bar{z_1} \cdot \bar{z_2}$.

内积具有如下性质：

(1) $(\theta,\ x) = (x,\ \theta) = 0$，　$\forall\, x \in X$.

因为$(\theta,\ x) = (0\theta,\ x) = 0(\theta,\ x) = 0$.

(2) $\forall\, z \in X$，当$(x,\ z) = (y,\ z)$ 时 $x = y$.

因为$(x,\ z) - (y,\ z) = 0$，故$(x - y,\ z) = 0$. 由 z 的任意性，取 $z = x - y$，则有$(x - y,\ x - y) = 0$，从而 $x - y = \theta$，即 $x = y$.

(3) $(z,\ \alpha x + \beta y) = \bar{\alpha}(z,\ x) + \bar{\beta}(z,\ y)$.

因为　　　　$(z,\ \alpha x + \beta y) = \overline{(\alpha x + \beta y,\ z)}$

$$= \overline{\alpha(x,\ z)} + \overline{\beta(y,\ z)}$$
$$= \bar{\alpha}\ \overline{(x,\ z)} + \bar{\beta}\ \overline{(y,\ z)}$$
$$= \bar{\alpha}(z,\ x) + \bar{\beta}(z,\ y).$$

(4)当 X 是实内积空间时,内积公理的 2°, 3°为:

(2′) $(x,\ y) = \overline{(y,\ x)} = (y,\ x)$;

(3′) $(z,\ \alpha x + y\beta) = \alpha(z,\ x) + \beta(z,\ y)$.

例 1 (1)在 $\mathbf{E}^n$ 中, $\forall x = (\xi_1,\ \xi_2,\ \cdots,\ \xi_n)$, $y = (\eta_1,\ \eta_2,\ \cdots,\ \eta_n)$. 定义内积

$$(x,\ y) = \sum_{i=1}^{n} \xi_i \eta_i.$$

(2)在 $\mathbf{C}^n$ 中定义内积

$$(x,\ y) = \sum_{i=1}^{n} \xi_i\ \overline{\eta_i}.$$

容易验证,以上定义满足内积公理而分别成为实和复内积空间.

定理 1 (Cauchy-Schwarz 不等式)设 X 是内积空间,则有

$$|(x,\ y)| \leqslant \sqrt{(x,\ x)(y,\ y)},\ \forall x,\ y \in X.$$

证 (1)当 $y = \theta$ 时不等式显然成立.

(2) 当 $y \neq \theta$ 时, $\forall \alpha \in K$ 有

$$(x - \alpha y,\ x - \alpha y)$$
$$= (x,\ x - \alpha y) - \alpha(y,\ x - \alpha y)$$
$$= (x,\ x) - \bar{\alpha}(x,\ y) - \alpha(y,\ x) + \alpha\bar{\alpha}(y,\ y).$$

令 $\alpha = \dfrac{\overline{(y,\ x)}}{(y,\ y)}$ 代入上式有

$$(x - \alpha y,\ x - \alpha y)$$
$$= (x,\ x) - \frac{(y,\ x)}{(y,\ y)} \cdot (x,\ y) - \frac{\overline{(y,\ x)}}{(y,\ y)}(y,\ x) + \frac{\overline{(y,\ x)}}{(y,\ y)} \cdot \frac{(y,\ x)}{(y,\ y)}(y,\ y)$$
$$= (x,\ x) - \frac{|(y,\ x)|^2}{(y,\ y)} - \frac{|(y,\ x)|^2}{(y,\ y)} + \frac{|(y,\ x)|^2}{(y,y)}$$
$$= (x,\ x) - \frac{|(y,\ x)|^2}{(y,\ y)}.$$

由 $(x - \alpha y,\ x - \alpha y) \geqslant 0$ 知 $(x,\ x)(y,\ y) \geqslant |(y,\ x)|^2 = |(x,\ y)|^2$, 于是

$$|(x, y)| \leqslant \sqrt{(x, x)(y, y)}.$$

二、内积空间与赋范线性空间关系

定理 2 设$(X, (\cdot))$是内积空间,定义 $\|x\| = \sqrt{(x, x)}, \forall x \in X$,则$(X, \|\cdot\|)$成为赋范空间.

证 由 $|(x, y)| \leqslant \sqrt{(x, x)(y, y)} = \|x\| \|y\|$ 可得

$$\begin{aligned}
\|x+y\|^2 &= (x+y, x+y) \\
&= (x, x) + (x, y) + (y, x) + (y, y) \\
&= \|x\|^2 + (x, y) + \overline{(x, y)} + \|y\|^2 \\
&= \|x\|^2 + 2\mathrm{Re}\{(x, y)\} + \|y\|^2 \\
&\leqslant \|x\|^2 + 2|(x, y)| + \|y\|^2 \\
&\leqslant \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,
\end{aligned}$$

故
$$\|x+y\| \leqslant \|x\| + \|y\|.$$

另外,由范数公理知

$$\|x\| \geqslant 0, \ \|x\| = 0 \Leftrightarrow x = \theta;$$

由 $\|\alpha x\| = \sqrt{(\alpha x, \alpha x)}$ 知

$$\|\alpha x\|^2 = (\alpha x, \alpha x) = \alpha\bar{\alpha}(x, x) = |\alpha|^2 \|x\|^2,$$

故
$$\|\alpha x\| = |\alpha| \ \|x\|.$$

所以$(X, \|\cdot\|)$是赋范空间.

定义 2 若$(X, (\cdot, \cdot))$是内积空间,则说 $\|x\| = \sqrt{(x, x)}$ 是由内积导出的范数,$(X, \|\cdot\|)$是由内积导出的赋范线性空间.

推论 1 (内积的连续性)设 $x_n \to x, y_n \to y (n \to \infty)$,则$(x_n, y_n) \to (x, y)(n \to \infty)$,即内积$(x, y)$是 x, y 的二元连续函数.

证 由 $x_n \to x$ 知$\exists M > 0$,使得$\|x_n\| \leqslant M$. 所以

$$\begin{aligned}
|(x_n, y_n) - (x, y)| &= |(x_n, y_n) - (x_n, y) + (x_n, y) - (x, y)| \\
&\leqslant |(x_n, y_n) - (x_n, y)| + |(x_n, y) - (x, y)| \\
&\leqslant |(x_n, y_n - y)| + |(x_n - x, y)| \\
&\leqslant \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \\
&\leqslant M\|y_n - y\| + \|x_n - x\| \|y\| \to 0.
\end{aligned}$$

即 $\lim\limits_{n\to\infty}(x_n, y_n) = (x, y)$.

当 X 是复空间时，

$$(x, y) = \frac{1}{4}[\|x+y\|^2 - \|x-y\|^2 + \mathrm{i}\|x+\mathrm{i}y\|^2 - \mathrm{i}\|x-\mathrm{i}y\|^2], \tag{1}$$

其中 $\mathrm{i} = \sqrt{-1}$；

当 X 是实内积空间时

$$(x, y) = \frac{1}{4}[\|x+y\|^2 - \|x-y\|^2]. \tag{2}$$

式(1)，式(2)称为复(实)内积空间中的极化恒等式，这两个式子说明可以通过范数来计算内积.

由定理 2 可知，内积空间$\xrightarrow{\text{导出}}$赋范空间$\xrightarrow{\text{导出}}$度量空间，故度量空间、赋范空间的概念和结论在内积空间中也适用.

定义 3　设 X 是内积空间，若导出的$(X, \|\cdot\|)$是完备的，则说 X 是 Hilbert空间.

例 2　$\mathbf{C}^n$ 中，定义 $(x, y) = \sum\limits_{i=1}^{n}\xi_i\bar{\eta}_i$，$\|x\| = \sqrt{(x, x)} = \left(\sum\limits_{i=1}^{n}|\xi_i|^2\right)^{\frac{1}{2}}$，由$(\mathbf{C}^n, \|\cdot\|)$是完备的可知$(\mathbf{C}^n, (\cdot, \cdot))$是Hilbert空间.

例 3　(1)有限维内积空间，l^2，$L^2[a, b]$都是Hilbert空间.

(2) 若 $C[a, b]$是$[a, b]$上复值连续函数全体，定义 $(x, y) = \int_a^b x(t)\overline{y(t)}\,\mathrm{d}t$，$x, y \in C[a, b]$，则 $C[a, b]$是内积空间，导出范数 $\|x\|_2 = \left(\int_a^b |x(t)|^2\mathrm{d}t\right)^{\frac{1}{2}}$.

由$(C[a, b], \|\cdot\|)$不完备可知 $C[a, b]$按以上定义的内积不是Hilbert空间.

注意　不是每一个内积都能由某个范数导出.

定理 2　赋范空间$(X, \|\cdot\|)$能成为内积空间的充分必要条件是 $\|\cdot\|$ 满足：

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2),\ \forall x, y \in X.$$

此公式称为平行四边形公式.

例 4　$(L^1[0, 2], \|\cdot\|_1)$不能成为内积空间，其中 $\|x\|_1 = \int_0^2 |x(t)|\,\mathrm{d}t$.

证　取 $x(t) \equiv 1$，$y(t) = t$，有

$$\| x+y \|_1 = \int_0^2 |1+t| \, dt = \int_0^2 (1+t) dt = 4,$$

$$\| x-y \|_1 = \int_0^2 |1-t| \, dt$$

$$= \int_0^1 (1-t) dt + \int_1^2 (t-1) dt = 1,$$

而 $$\| x \|_1 = \int_0^2 1 dt = 2, \quad \| y \|_1 = \int_0^2 |t| \, dt = 2,$$

故 $$\| x+y \|_1^2 + \| x-y \|_1^2 = 4^2 + 1 = 17$$

$$\neq 2(\| x \|_1^2 + \| y \|_1^2) = 8 \times 2 = 16.$$

同样 $L^p[a, b]$，$p \neq 2$ 时，都不是内积空间.

第二节　内积空间中的正交与投影

定义 1　设 X 是内积空间，$x, y \in X$，$M, N \subset X$，则

(1) 若 $(x, y) = 0$，则说 x 与 y 正交，记 $x \perp y$；

(2) 若 $\forall y \in M$，总有 $x \perp y$，则说 x 与 M 正交，记 $x \perp M$；

(3) 若 $\forall x \in M$，$\forall y \in N$，总有 $x \perp y$，则说 M 与 N 正交，记 $M \perp N$；

(4) $M^{\perp} = \{x \mid x \in X, x \perp M\}$ 叫做 M 在 X 中的正交补.

定理 1　（正交与正交补的性质）

(1) $x \perp y$ 时有 $\| x+y \|^2 = \| x \|^2 + \| y \|^2$；

(2) $M \cap M^{\perp} \subset \{\theta\}$，当 M 是 X 的子空间时 $M \cap M^{\perp} = \{\theta\}$；

(3) $M^{\perp}$ 是 X 的闭子空间；

(4) $\overline{M} = X$ 时 $M^{\perp} = \{\theta\}$.

证　(1) 由于 $x \perp y$，即 $(x, y) = (y, x) = 0$，故

$$\begin{aligned} \| x+y \|^2 &= (x+y, x+y) \\ &= (x, x) + (x, y) + (y, x) + (y, y) \\ &= \| x \|^2 + \| y \|^2. \end{aligned}$$

用归纳法易证：若 $\forall x_i \perp x_j$，$i \neq j$，则

$$\| \sum_{i=1}^{n} x_i \|^2 = \sum_{i=1}^{n} \| x_i \|^2.$$

注意　$\| x+y \|^2 = \| x \|^2 + \| y \|^2$ 不一定有 $x \perp y$；若 X 是实内积空间时，$x \perp y \Leftrightarrow \| x+y \|^2 = \| x \|^2 + \| y \|^2$；

(2) 若 $x \in M \cap M^{\perp}$,即 $x \in M, x \in M^{\perp}$,由 $x \in M^{\perp}$ 可知 $x \perp M$,故 $x \perp x$,所以 $(x, x) = 0$, $x = \theta$.

所以 $M \cap M^{\perp} \subset \{\theta\}$.

当 M 是 X 的子空间时,$M \cap M^{\perp} = \{\theta\}$,这是因为 $\theta \in M$.

(3) 先证 $M^{\perp}$ 是 X 的子空间:

设 $x, y \in M^{\perp}$, $\alpha, \beta \in K$, $\forall z \in M$ 有

$$(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) = 0,$$

故 $\alpha x + \beta y \in M^{\perp}$,所以 $M^{\perp}$ 为 X 的子空间.

再证 $M^{\perp}$ 是闭的:

若 $x_n \in M^{\perp}$,且 $x_n \to x_0$,则由内积的连续性知,对 $\forall y \in M$ 有

$$(x_0, y) = \lim_{n\to\infty}(x_n, y) = 0,$$

故 $x_0 \in M^{\perp}$. 所以 $M^{\perp}$ 是闭的线性子空间.

(4) 设 $\overline{M} = X$, $\forall x_0 \in M^{\perp} \subset X$,则 $\forall y \in M$,有 $(x_0, y) = 0$. 由 $x_0 \in \overline{M}$ 知 $\exists x_n \in M$,使得 $x_n \to x_0$,故由内积连续性知

$$(x_0, x_0) = \lim_{n\to\infty}(x_n, x_0) = 0 \Rightarrow x_0 = \theta.$$

所以
$$M^{\perp} = \{\theta\}.$$

定义 2 **X 是内积空间,M 是 X 的线性子空间,$x \in X$,若 $\exists x_0 \in M$,使得 $x - x_0 \in M^{\perp}$,则 x_0 叫做 x 在 M 上的正交投影,简称投影,表示式 $x = x_0 + (x - x_0)$ 叫做 x 关于 M 的正交分解.**

x 在 M 上的投影未必存在.

引理 1 **(正交分解的惟一性)设 X 为内积空间,M 是 X 的子空间,$x \in X$,若 x 在 M 上的正交投影存在,则投影惟一.**

证 设 x_0, x_1 是 x 在 M 上的投影,由定义

$$x_0, x_1 \in M \quad 且 \quad x - x_0, x - x_1 \in M^{\perp}.$$

因为 M, $M^{\perp}$ 都是 X 的子空间,所以 $x_0 - x_1 \in M$,且

$$(x - x_1) - (x - x_0) = x_0 - x_1 \in M^{\perp}.$$

由定理 1(2)知

$$x_0 - x_1 \in M \cap M^{\perp} \subset \{\theta\}$$

所以
$$x_0 - x_1 = \theta, \ x_0 = x_1.$$

引理 2 **设 X 是内积空间,M 是 X 的子空间,$x \in X$, $x_0 \in M$,则 x_0 为 x 在**

M **上的投影的充分必要条件是** $\|x-x_0\|=d(x, M)$.

证 必要性:设 x_0 是 x 在 M 上的投影,则$x_0\in M$且 $x-x_0\in M^{\perp}$. 所以 $\forall y\in M$,有 $x-x_0\perp y$. 又 $x-x_0\perp x_0$,所以 $x-x_0\perp x_0-y\in M$. 由 $x-y=(x-x_0)+(x_0-y)$ 及定理 1(1)有

$$\|x-y\|^2=\|x-x_0\|^2+\|x_0-y\|^2\geqslant\|x-x_0\|^2,$$

$$\|x-x_0\|\leqslant\inf_{y\in M}\|x-y\|=d(x, M). \tag{1}$$

若 $x\notin M$,则 $\|x-x_0\|\geqslant d(x, M)$,从而 $\|x-x_0\|=d(x, M)$.

若 $x\in M$,由式(1) 知 $\|x-x_0\|=d(x, M)=0$.

充分性证略.

定理 2 (**投影定理**)X **是内积空间**,M **是** X **的完备子空间,则** $\forall x\in X$, x **在** M **上的投影都存在.**

证 设 $d=d(x, M)$,则 $\exists\{x_n\}\subset M$,有 $\|x_n-x\|\to d$.

下证$\{x_n\}$是 M 中的 Cauchy 列:

由平行四边形公式有

$$\|x\|^2+\|y\|^2=2\left(\left\|\frac{x+y}{2}\right\|^2+\left\|\frac{x-y}{2}\right\|^2\right),$$

令
$$u=x_m-x,\qquad v=x_n-x,$$

则
$$\frac{u+v}{2}=\frac{x_m+x_n}{2}-x,$$

$$\frac{u-v}{2}=\frac{x_m-x_n}{2}.$$

于是
$$2\left(\left\|\frac{u+v}{2}\right\|^2+\left\|\frac{u-v}{2}\right\|^2\right)=\|u\|^2+\|v\|^2,$$

$$2\left\|\frac{x_m-x_n}{2}\right\|^2=\|x_m-x\|^2+\|x_n-x\|^2-2\left\|\frac{x_m+x_n}{2}-x\right\|^2.$$

由
$$\frac{x_m+x_n}{2}\in M \text{ 知 } \left\|\frac{x_m+x_n}{2}-x\right\|\geqslant d,$$

所以
$$0\leqslant 2\left\|\frac{x_m-x_n}{2}\right\|^2\leqslant\|x_m-x\|^2+\|x_n-x\|^2-2d^2.$$

分别令 $m, n\to\infty$,由 $\|x_n-x\|\to d$, $\|x_m-x\|\to d$ 可得

$$\left\|\frac{x_m-x_n}{2}\right\|^2\to 0.$$

故 $\| x_m - x_n \| \to 0$，即$\{x_n\}$是M中的 Cauchy 列.

由M是完备知 $\exists x_0 \in M$,使得 $x_n \to x_0$. 而 $\| \cdot \|$ 连续知 $d = \lim\limits_{n\to\infty} \| x - x_n \| = \| x - x_0 \|$. 由引理 2 知 x_0 是 x 在M上的投影，而且惟一.

推论 1 若 X 是 Hilbert 空间，M 是 X 的闭子空间的充分必要条件是 $\forall x \in X$，x 在M中的投影 x_0 都存在，且

$$\| x - x_0 \| = d(x, M).$$

由定理 2 和引理 2 可得

定理 3 设 X 是 Hilbert 空间，M 是 X 的子空间，则 $\forall x \in X$，x 在M上投影存在的充分必要条件是M是 X 的闭子空间.

推论 2 若 X 是 Hilbert 空间，M 是 X 的真闭子空间，则 $M^{\perp} \neq \{\theta\}$.

证 取 $x \in X - M$. 由推论 1 知，$\exists x_0 \in M$，x_0 是 x 的投影，故 $x - x_0 \in M^{\perp}$，使$x = x_0 + (x - x_0)$.

又 $x \notin M$，故 $x \neq x_0$，所以 $x - x_0 \neq \theta$.

推论 3 设 X 是 Hilbert 空间，M 是 X 的闭子空间，则 $M = (M^{\perp})^{\perp}$.

证 $\forall x \in M$，则 $x \perp M^{\perp}$，所以$x \in (M^{\perp})^{\perp}$，即$M \subset (M^{\perp})^{\perp}$.

又因$(M^{\perp})^{\perp}$是 $M^{\perp}$ 的正交补，所以$(M^{\perp})^{\perp}$为 X 的闭子空间.

任取 $x \in (M^{\perp})^{\perp} \subset X$，关于 M 作正交分解$x = x_0 + x_1$，其中 $x_0 \in M$，$x_1 \in M^{\perp}$. 注意到

$$x,\ x_0 \in (M^{\perp})^{\perp},$$

所以
$$x_1 = x - x_0 \in (M^{\perp})^{\perp},$$

于是
$$x_1 \in M^{\perp} \cap (M^{\perp})^{\perp} = \{\theta\},$$

即
$$x = x_0 \in M,$$

从而
$$M = (M^{\perp})^{\perp}.$$

一般讲，当 M 为 Hilbert 空间 X 的子空间时，有 $\overline{M} = (M^{\perp})^{\perp}$.

内积空间本身也有正交分解概念.

定义 3 设 X 是内积空间，M_1，M_2 是 X 的子空间，$M_1 \perp M_2$. 若 $\forall x \in X$，都有$x = x_1 + x_2$，其中$x_1 \in M_1$，$x_2 \in M_2$，则说 X 是 M_1 和 M_2 的正交和，记为 $X = M_1 \oplus M_2$

注意 若 $X = M_1 \oplus M_2$，则 $M_1 \cap M_2 = \{\theta\}$，并且 $M_1 = M_2^{\perp}$，$M_2 = M_1^{\perp}$.

投影定理也可表述成如下空间正交分解的形式.

定理 4 X 是内积空间，M 是 X 的完备子空间，则 $X = M \oplus M^{\perp}$.

第三节 内积空间中的 Fourier 级数

在 $\mathbf{E}^3$ 中，有三个相互正交的单位向量 $e_1=\{1,0,0\}$，$e_2=\{0,1,0\}$，$e_3=\{0,0,1\}$，有了这三个向量，则任何向量 a 有惟一的分解式

$$a=a_1e_1+a_2e_2+a_3e_3, \tag{1}$$

其中，$a_1=(a,e_1)$，$a_2=(a,e_2)$，$a_3=(a,e_3)$. 式(1) 用到了向量的正交性. 在内积空间中，我们有了正交的概念，设法把内积空间中的元素表示成式(1)的表达式，将会是十分有意义的.

一、标准正交系

定义 1 设 X 是内积空间，$\{e_n\}$ 是 X 中的点列，若满足：

$$(e_i,e_j)=\begin{cases}0, & \text{当 } i\neq j \text{ 时};\\ 1, & \text{当 } i=j \text{ 时},\end{cases}$$

则说 $\{e_n\}$ 是 X 中的一个标准正交系.

例如，$e_1=(1,0,\cdots,0,\cdots)$，$e_2=(0,1,\cdots,0,\cdots)$，$e_k=(0,0,\cdots,1,0,\cdots)\cdots$ 是 l^2 中的标准正交系.

又如

$$\left\{\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos x,\ \frac{1}{\sqrt{\pi}}\sin x,\cdots,\frac{1}{\sqrt{\pi}}\cos nx,\ \frac{1}{\sqrt{\pi}}\sin x,\ \cdots\right\}$$

是 $L^2[-\pi,\pi]$中的标准正交系.

标准正交系有如下简单性质：

标准正交系 $\{e_n\}$ 是线性独立系，即 $\{e_n\}$ 中任何有限组是线性无关的.

事实上设 $\{e_{n_1},e_{n_2},\cdots,e_{n_k}\}$ 是 $\{e_n\}$ 中的任一有限组，如果

$$a_1e_{n_1}+a_2e_{n_2}+\cdots+a_ke_{n_k}=\theta,$$

则对 $\forall j=1,2,\cdots,k$，有

$$0=(\theta,e_{n_j})=\Big(\sum_{i=1}^{n}a_ie_{n_i},\ e_{n_j}\Big)=a_j(e_{n_j},\ e_{n_j})=a_j,$$

故 $\{e_{n_1},e_{n_2},\cdots,e_{n_k}\}$ 线性无关.

定理 1 （正交化定理）**设 $\{x_n\}$ 是内积空间 X 中的任一线性独立系，则可将 $\{x_n\}$ 标准正交化为标准正交系 $\{e_n\}$，且对任何正整数 n，有**

$$x_n = \sum_{k=1}^{n} a_k^{(n)} e_k, \qquad e_n = \sum_{k=1}^{n} \beta_k^{(n)} x_k$$

且 $\mathrm{span}\{e_1, e_2, \cdots, e_n\} = \mathrm{span}\{x_1, x_2, \cdots, x_n\}$.

证 因$\{x_n\}$是线性独立系,所以 $x_1 \neq 0$.

令 $e_1 = \dfrac{x_1}{\| x_1 \|}$, 则 $\| e_1 \| = 1$.

再将 x_2 在由 e_1 所张成的子空间 M_1 上作正交分解

$$x_2 = (x_2, e_1)e_1 + v_2,$$

其中 $v_2 = x_2 - (x_2, e_1)e_1$. 由于 x_1, x_2 线性无关,因而 $v_2 \neq 0$ 且$v_2 \in M_1^{\perp}$ (事实上,

$$(v_2, e_1) = (x_2, e_1) - (x_2, e_1)(e_1, e_1) = 0).$$

即 $v_2 \perp e_1$,再令 $e_2 = \dfrac{v_2}{\| v_2 \|}$,则 $\| e_2 \| = 1$,显见 e_2 是 x_1, x_2 的线性组合,且 $e_2 \perp e_1$.

再将 x_3 在由 e_1, e_2 所张成的子空间 M_2 上作正交分解,

$$x_3 = (x_3, e_1)e_1 + (x_3, e_2)e_2 + v_3,$$

易知 $v_3 = x_3 - (x_3, e_1)e_1 - (x_3, e_2)e_2 \neq 0$,且 $v_3 \in M_2^{\perp}$. 令 $e_3 = \dfrac{v_3}{\| v_3 \|}$, $\| e_3 \| = 1$,则 e_3 是 x_1, x_2, x_3 的线性组合且$e_3 \perp e_2$, $e_3 \perp e_1$.

应用数学归纳法可以证明,对每一个 n,

$$v_n = x_n - \sum_{i=1}^{n-1} (x_n, e_i)e_i$$

不是零向量且 v_n 与 e_1, e_2, $\cdots$, e_{n-1}都正交,取 $e_n = \dfrac{v_n}{\| v_n \|}$, $\| e_n \| = 1$.

显然

$$x_n = v_n + \sum_{i=1}^{n-1} (x_n, e_i)e_i$$

$$= \sum_{i=1}^{n-1} (x_n, e_i)e_i + \| v_n \| e_n = \sum_{i=1}^{n} a_i^{(n)} e_i,$$

反之容易看出,e_n 是 x_1, x_2, $\cdots$, x_n 的线性组合. 从而 $\mathrm{span}\{e_1, e_2, \cdots, e_n\} = \mathrm{span}\{x_1, x_2, \cdots, x_n\}$.

二、Fourier 级数及其收敛性

定义 2 设 X 是内积空间,$\{e_n\}$是 X 中的标准正交系, $x \in X$, 则说内积

$C_k = (x, e_k)(k = 1, 2, 3, \cdots)$ 是 x 关于 $\{e_n\}$ 的Fourier 系数，

$$\sum_{k=1}^{\infty} C_k e_k = \sum_{k=1}^{\infty} (x, e_k) e_k$$

叫做 x 关于 $\{e_n\}$ 的Fourier 级数.

定理 2 设 X 是内积空间，$\{e_n\}$ 是 X 的标准正交系，取定 $e_{n_1}, e_{n_2}, \cdots, e_{n_k} \in \{e_n\}$，记 $M = \operatorname{span}\{e_{n_1}, e_{n_2}, \cdots, e_{n_k}\}$，$\forall x \in X$，则

$$x_0 = \sum_{i=1}^{k} (x, e_{n_i}) e_{n_i}$$

是 x 在 M 上的正交投影.

证 由于 M 是 X 的 k 维子空间，故 M 是 X 的闭的线性子空间，亦即 M 作为内积空间是完备的，由投影定理知 x 在 M 中存在惟一的正交投影. 下面验证 x_0 是 x 在 M 中的正交投影. 因为

$$(x_0, e_{n_j}) = \sum_{i=1}^{k} (x, e_{n_i})(e_{n_i}, e_{n_j}) = (x, e_{n_j}), \ j = 1, 2, 3, \cdots, k$$

故有
$$(x - x_0, e_{n_i}) = (x, e_{n_i}) - (x_0, e_{n_i}) = 0.$$

$\forall y \in M$，则 $y = \sum_{i=1}^{k} \alpha_i e_{n_i}$，$\alpha_i \in \Lambda$，且有

$$(x - x_0, y) = \sum_{i=1}^{k} \bar{\alpha}_i (x - x_0, e_{n_i}) = 0.$$

故 $(x - x_0) \perp M$，所以 x_0 是 x 在 M 上的正交投影.

定理 3 （最值逼近定理）设 $\{e_n\}$ 是内积空间 X 的标准正交系，$x \in X$，$\{C_k\} = \{(x, e_k)\}$，$k = 1, 2, \cdots, n$，则对任何数组 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，$\alpha_i \in \Lambda$，有

$$\| x - \sum_{k=1}^{n} C_k e_k \| \leqslant \| x - \sum_{k=1}^{n} \alpha_k e_k \|.$$

证 设 $x_n = \sum_{k=1}^{n} C_k e_k$，由本节定理 2 知，$x_n$ 是 x 在 $M = \operatorname{span}\{e_1, e_2, \cdots, e_n\}$ 中的正交投影，则 $x - x_n = x - \sum_{k=1}^{n} C_k e_k \in M^{\perp}$. 设

$$x_n' = \sum_{k=1}^{n} \alpha_k e_k \in M,$$

故
$$x_n - x_n' = \sum_{k=1}^{n} (C_k - \alpha_k) e_k \in M,$$

应用勾股定理有

$$\begin{aligned}\| x-\sum_{k=1}^{n}\alpha_k e_k \|^2 &= \| (x-x_n)-(x_n-x'_n) \|^2 \\ &= \| x-x_n \|^2+\| x_n-x'_n \|^2 \\ &= \| x-\sum_{k=1}^{n}C_k e_k \|^2+\| x_n-x'_n \|^2 \\ &\geqslant \| x-\sum_{k=1}^{n}C_k e_k \|^2.\end{aligned}$$

注意 本定理说明，当用 n 个标准正交向量的线性组合来逼近 X 的向量 x 时，只有取系数为 Fourier 系数，逼近程度最好.

现在我们研究向量 x 的 Fourier 级数是否一定收敛？如果收敛，是否一定收敛到 x？

定理 4 (Bessel(贝塞尔)不等式)**设** $\{e_n\}$ **为内积空间** X **中标准正交系，** $x\in X$，**则** $\sum_{i=1}^{\infty}|(x, e_n)|^2$ **收敛，且** $\sum_{n=1}^{\infty}|(x, e_n)|^2\leqslant \| x \|^2$.

证 令 $C_k=(x, e_k)$，$x_n=\sum_{k=1}^{n}C_k e_k$，由本节定理 2 知，$(x-x_n)\perp x_n$，由勾股定理有

$$\| x \|^2=\| (x-x_n)+x_n \|^2=\| x-x_n \|^2+\| x_n \|^2\geqslant \| x_n \|^2,$$

$$\begin{aligned}\| x_n \|^2 &= (x_n, x_n)=\left(\sum_{j=1}^{n}C_j e_j, \sum_{k=1}^{n}C_k e_k\right) \\ &= \sum_{k=1}^{n}\left(\sum_{j=1}^{n}C_j e_j, C_k e_k\right) \\ &= \sum_{k=1}^{n}\sum_{j=1}^{n}C_j\overline{C_k}(e_j, e_k)=\sum_{k=1}^{n}C_k\overline{C_k}=\sum_{k=1}^{n}|C_k|^2.\end{aligned}$$

于是有 $\quad \sum_{k=1}^{n}|C_k|^2\leqslant \| x \|^2.$

令 $n\to\infty$，则有 $\sum_{k=1}^{\infty}|C_k|^2\leqslant \| x \|^2$.

注意 由定理证明，得到一个常用公式

$$\| x_n \|^2=\sum_{k=1}^{n}|(x, e_k)|^2 \tag{1}$$

定理 5 设$\{e_n\}$是内积空间 X 的标准正交系，$x\in X$，则 x 关于$\{e_n\}$的 Fourier 级数收敛于x的充分必要条件是

$$\|x\|^2=\sum_{k=1}^{\infty}|(x,e_k)|^2 \tag{2}$$

成立，式(2)叫做 Parseval(帕塞瓦尔)公式.

证 取 $x_n=\sum_{k=1}^{n}(x,e_k)e_k$，则 $x-x_n\perp x_n$，由勾股定理及式(1)可得

$$\|x-\sum_{k=1}^{n}(x,e_k)e_k\|^2=\|x-x_n\|^2=\|x\|^2-\|x_n\|^2$$

$$=\|x\|^2-\sum_{k=1}^{n}|(x,e_k)|^2,$$

从而

$$\lim_{n\to\infty}x_n=x\Leftrightarrow\lim_{n\to\infty}\|x-\sum_{k=1}^{n}(x,e_k)e_k\|=0$$

$$\Leftrightarrow\|x\|^2=\sum_{k=1}^{\infty}|(x,e_k)|^2.$$

定义 3 说内积空间 X 的标准正交系$\{e_n\}$是完全的，如果 $x\in X$，当 $x\perp e_n$，$n=1,2,3,\cdots$，则有$x=\theta$.

定理 6 设 X 是 Hilbert 空间，则以下命题等价：

(1) $\{e_n\}$是 X 中的完全标准正交系；

(2) $\forall x\in X$，有 $\sum_{k=1}^{\infty}(x,e_k)e_k=x$；

(3) $\forall x\in X$，Parseval 公式成立，即

$$\|x\|^2=\sum_{k=1}^{\infty}|(x,e_k)|^2.$$

证 (1)⇒(2)：设$\{e_n\}$是 Hilbert 空间 X 中的完全标准正交系，令 $M=\operatorname{span}\{e_n\}$，易证 $\overline{M}=X$. 事实上，若不然，则存在$x\in X-\overline{M}$，$x\neq\theta$，且有分解式$x=x_0+z$，$x_0\in\overline{M}$，$z\in\overline{M}^{\perp}$. 显然$z\neq\theta$，则$z\perp e_n$，$n=1,2,3,\cdots$，这与$\{e_n\}$完全相矛盾，即 M 在 X 中稠密.

由此 $\forall x\in X$，$\forall\varepsilon>0$，存在$y'=\sum_{k=1}^{N}\alpha_k e_k\in M$，使得

$$\|x-y'\|=\|x-\sum_{k=1}^{N}\alpha_k e_k\|<\varepsilon.$$

令 $C_k=(x,e_k)$，由本节定理 3 有

$$\| x-\sum_{k=1}^{N}C_ke_k \| \leqslant \| x-\sum_{k=1}^{N}\alpha_ke_k \| <\varepsilon,$$

故当 $m>N$ 时，由勾股定理有

$$\| x-\sum_{k=1}^{m}C_ke_k \|^2 = \| x\|^2-\sum_{k=1}^{m}|C_k|^2$$

$$\leqslant \| x\|^2-\sum_{k=1}^{N}|C_k|^2 = \| x-\sum_{k=1}^{N}C_ke_k \|^2<\varepsilon^2,$$

令　$m\to+\infty$ 有　$\| x-\sum_{k=1}^{\infty}C_ke_k \|^2\leqslant\varepsilon^2.$

由 $\varepsilon>0$ 任意性知　　$x=\sum_{k=1}^{\infty}C_ke_k.$

(2) ⇒(3)：由定理 5 可得.

(3) ⇒(1)：设 $\forall x\in X$，有 $\| x\|^2=\sum_{k=1}^{\infty}|C_k|^2$，如果 $x\perp e_n$，$n=1,2,3,\cdots$，则 $C_k=(x,e_k)=0$.

故 $\| x\|^2=\sum_{k=1}^{\infty}|C_k|^2=0$，从而 $x=\theta$.

即 $\{e_n\}$ 是 X 中的完全标准正交系.

到目前为止，我们讲到了许多空间，把它们之间的关系梳理如下图（符号 ∪ 表示上面的空间包含着下面的空间）.

拓扑空间		线性空间		
∪		∪		
Hausdorff 空间		线性拓扑空间		
∪		∪		
度量空间	⊃	线性度量空间	⊃	完备线性度量空间
		∪		∪
		赋范线性空间	⊃	Banach 空间
		∪		∪
		内积空间	⊃	Hilbert 空间
				∪
				欧氏空间 $\mathbf{E}^n$ 和 $\mathbf{C}^n$

习　题　3-2

1. 设 X 为内积空间，$x, y \in X$，若对一切 $z \in X$ 均有 $(x, z) = (y, z)$，试证：$x = y$.

2. 设 X 是内积空间，$x_n, x \in X$ 且 $\|x_n\| \to \|x\|$ 及 $(x_n, x) \to (x, x)$，试证：$x_n \xrightarrow{\|\cdot\|} x$.

3. 设 H 是内积空间，$x_1, x_2, \cdots, x_n$ 是 H 中的向量，且满足条件

$$(x_\mu, x_\nu) = \begin{cases} 0, & \text{当 } \mu \neq \nu \text{ 时}; \\ 1, & \text{当 } \mu = \nu \text{ 时}. \end{cases}$$

试证：$\{x_1, x_2, \cdots, x_n\}$ 是线性无关的.

4. 设 X 是实内积空间，x, y 为非零元，证明：

$$\|x+y\| = \|x\| + \|y\| \Leftrightarrow \exists \lambda > 0 \text{ 使得 } y = \lambda x.$$

5. 设 X 为内积空间，$y \in X$，试证：映射 $f_y: X \to K$ 且 $f_y(x) = (x, y)$ 是 X 上线性连续泛函，且 $\|f_y\| = \|y\|$.

6. 证明：$C[0, 1]$，$l^p\ (p \neq 2)$ 不是内积空间.

7. 若 X 为实内积空间，试证：

$$x \perp y \Leftrightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

又若 X 为复内积空间，上面结论是否成立？

8. 若 X 为实内积空间，试证：

$$(x+y) \perp (x-y) \Leftrightarrow \|x\| = \|y\|.$$

当 $X = \mathbf{E}^2$ 时，说明上面表示式的几何意义. 又若 X 为复内积空间，条件 $\|x\| = \|y\|$ 意味着什么？

9. 设 X 为内积空间，K 为数域，试证：

$$x \perp y \Leftrightarrow \|x+\lambda y\| = \|x-\lambda y\|,\ \forall \lambda \in K \text{ 成立}.$$

10. 设 X 为内积空间，K 为数域，试证：

$$x \perp y \Leftrightarrow \|x+\lambda y\| \geqslant \|x\|,\ \forall \lambda \in K.$$

11. 设 X 是内积空间，M 和 N 是 X 的子集，证明：当 $M \perp N$ 时，$M \subset N^\perp$，$N \subset M^\perp$.

12. 设 M 为Hilbert空间 H 的子空间，若 $\forall x \in H$，x 在 M 上投影 x_0 均存在，则 M 为 H 的闭子空间.

13. 设 M 是Hilbert空间 H 的非空子集，试证：$(M^\perp)^\perp$ 是 H 中包含 M 的最小闭子空间.

14. 设 $\{e_n\}_{n=1}^\infty$ 是Hilbert空间 H 中标准正交列，证明：$\{e_n\}_{n=1}^\infty$ 是标准正交基的充分必要条件是 $(x, y) = \sum\limits_{n=1}^\infty (x, e_n)\overline{(y, e_n)}$，$\forall x, y \in H$.

15. 设 $F = \{e_n\}$ 为Hilbert空间 H 中标准正交列，证明：数项级数 $\sum\limits_{n=1}^\infty |a_n|^2$ 收敛的充分必要条件是 $\sum\limits_{n=1}^\infty a_n e_n$ 依 H 中范数收敛.

参 考 文 献

1. 夏道行等.实变函数与泛函分析.北京:高等教育出版社,1985
2. 胡适耕.泛函分析.北京:高等教育出版社,2001
3. 张恭庆等.泛函分析讲义(上册).北京:北京大学出版社,1987
4. 李大华.应用泛函简明教程.武汉:华中理工大学出版社,1999
5. 许天明.应用泛函分析.北京:科学出版社,2002
6. 李广民等.应用泛函分析原理.西安:西安电子科技大学出版社,2003
7. 杜珣.现代数学引论.北京:北京大学出版社,1996

第四篇 抽 象 代 数

抽象代数是现代数学的一个重要分支.抽象代数从它产生的年代就明显有别于古典代数学,它的主要研究对象不是代数结构中的元素特性,而是各种代数结构本身以及不同代数结构之间的相互联系.掌握抽象代数的内容和它所体现的丰富的数学思想和方法是非常重要的,它的内容和方法不仅深入到数学的各个分支,也深入到其他一些学科.学习抽象代数既可以了解这一学科的特点,也有助于培养严格的逻辑思维能力.随着科学技术的发展,抽象代数在通信理论、计算机科学、系统工程 、现代物理学、现代化学以及密码学等许多领域中有着广泛的应用,是研究这些科学不可或缺的工具.

本篇介绍抽象代数的基本内容:群、环、域等代数对象及基本性质.

第一章 群

群是抽象代数中最早的而且是最基本的一个代数系,它也是现代数学的一个极其重要的概念.本章简单地介绍群的基本概念,主要是说明群、子群、同构、同态、正规子群、商群等的意义,以及它们的基本性质.

第一节 群 的 概 念

一、群的定义

在数学各部分以及数学的应用中,很多代数系是只有一种结合法的.在只有一种结合法的代数系中,最重要的就是群,它的运算法则与数的运算法则类似,并且有非常广泛的应用,是抽象代数中最基本的代数系.

定义 1 一个非空集合 G,若满足下面 4 个条件,就叫做群:

(1) G 有一个闭合的结合法,也就是说,G 中任意两元 a, b 的结合 c 仍然是 G 中元.G 的这种结合法通常叫做乘法,这时 c 又叫做 a, b 的积,我们经常用记号

$$a \cdot b = c \quad 或 \quad ab = c$$

表示.要注意的是积 ab 虽然是由 a, b 惟一决定的,但它一般还与 a, b 的顺序有关,也即 ab 不一定等于 ba.

(2) 结合法对于结合律成立,即对于 G 中任意三个元 a, b, c,有

$$(ab)c = a(bc).$$

(3) G 中有一个(左)单位元 e,对于 G 中任意元 a,有

$$ea = a.$$

(4) 对于 G 中的每一个元 a,在 G 中存在一个(左)逆元 a^{-1},使得 $a^{-1}a = e$ 成立.

非空集合 G,若它满足上面(1),(2)两个条件,叫做半群,半群也是一个重要的代数系.当 G 只满足(1)时,叫做乘集.

定义 2　一个群,假如它的结合法还满足交换律:

$$a \cdot b = b \cdot a.$$

这群就叫做交换群或Abel(阿贝尔)群,否则叫做非交换群.

为了对群这个概念有较深入的认识,下面给出一些例.

例 1　G 只包含一个元 g,乘法是 $gg = g$, G 对于这个乘法来说成为一个群,叫做单位元群.

例 2　G 是全体整数的集合 $\mathbf{Z}$, G 的乘法是通常整数的普通加法,则 G 对于加法成为一个群,叫做整数加群.常把积ab写成 $a+b$,单位元是 0,元 a 的逆元写成 $-a$.

例 3　G 是全体正有理数的集合 $\mathbf{Q}^+$,G 的结合法是通常的普通乘法,那么 G 成为一个群,它的单位元为 1.全体有理数的集合 $\mathbf{Q}$ 对于加法也成为群,单位元为 0,但对于乘法不能成为群,因为零没有逆元.

例 4　设 $G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$,则 G 对于矩阵的乘法成为一个群,单位元是 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

例 5　设 $M = \{(a, b) \mid a, b \in \mathbf{R}, a \neq 0\}$,其中 $\mathbf{R}$ 是实数集.若在 M 中规定,当 (a, b), $(c, d) \in M$,则有 $(a, b)(c, d) = (ac, bc + d)$,那么 M 成群,$(1, 0)$是单位元,(a, b)的逆元是 $(a^{-1}, -ba^{-1})$.

定义 3　一个群的元的个数是一个有限整数时,这个群叫做有限群,否则叫做无限群.一个有限群 G 的元的个数叫做这个群的阶,用 $|G|$ 表示.

例如,上面例 1 是有限群,例 2 与例 3 的群是无限群.

二、群的基本性质

上面给出了群的一些例子，现在再从群的定义出发来讨论群的基本性质.

从群定义中条件(2)，可以知道任意三个元a，b，c的乘积，由它们自身及它们的顺序惟一决定，与结合的先后也就是与所加的"括弧"无关，对于群G的任意n个元的积也是如此.

定理 1 **群G中任意n个元a_1，a_2，…，a_n的乘积由它们自身及它们的顺序惟一决定.**

证 用归纳法证明.当$n=3$时，就是群定义中的(2)，现在假定元数小于n时定理成立，来证明元数是n时定理依然成立.

n个元依a_1，a_2，…，a_n的顺序的乘积不外为下列各种形式：

$$a_1\cdot(a_2\cdot a_3\cdots a_n),$$
$$(a_1\cdot a_2)\cdot(a_3a_4\cdots a_n),$$
$$\cdots,$$
$$(a_1\cdot a_2\cdots a_{n-1})\cdot(a_n).$$

但其中任意一种

$$\begin{aligned}(a_1\cdot a_2\cdots a_m)(a_{m+1}\cdots a_n)&=(a_1\cdot(a_2\cdots a_m))\cdot(a_{m+1}\cdots a_n)\\&=(a_1)\cdot((a_2\cdots a_m)\cdot(a_{m+1}\cdots a_n))\\&=(a_1)\cdot(a_2\cdots a_n)\\&=a_1\cdot(a_2\cdots a_n).\end{aligned}$$

这即是说，它们都与第一种一致，所以定理成立.

假如G是交换群，那么a_1，a_2，…，a_n的乘积由它们自身惟一决定，与它们的顺序无关.

以后为了使用及书写方便，a_1，a_2，…，a_n的乘积就用$a_1a_2\cdots a_n$表示，不另加括弧，而且又有：

$$\underbrace{a\cdots a}_{n\text{个}}=a^n,\quad \underbrace{a^{-1}\cdots a^{-1}}_{n\text{个}}=a^{-n},\quad a^0=e,\quad a^1=a.$$

定理 2 **群中左逆元又是右逆元，左单位元又是右单位元.而且群的单位元惟一，逆元惟一.**

证 设a为群G的元，a^{-1}为a的左逆元，e为左单位元.

因为$a^{-1}aa^{-1}=ea^{-1}=a^{-1}$，用$a^{-1}$的左逆元左乘得到

$$(a^{-1})^{-1}(a^{-1}aa^{-1})=(a^{-1})^{-1}a^{-1}$$

所以

$$eaa^{-1} = e, \qquad 即 \qquad aa^{-1} = e.$$

这就是说,a 的左逆元同时又是右逆元,因此 a^{-1} 叫做 a 的逆元.

又因为 $\qquad ae = aa^{-1}a = ea = a, \qquad$ 即 $\qquad ae = a,$

这也就是说,左单位元又是右单位元,因此 e 叫做单位元.

假设 e_1, e_2 分别为群 G 的两个单位元,那么

$$e_2 = e_1e_2 = e_1,$$

所以单位元惟一.

又假设 a_1, a_2 都是 a 的逆元,那么 $\qquad a_1a = e, a_2a = e.$

因为 $aa_1 = e$, 所以

$$a_2 = a_2(aa_1) = (a_2a)a_1 = ea_1 = a_1,$$

从而逆元惟一,定理得证.

由 $aa^{-1} = e$, 得到 a 是 a^{-1} 的逆元,所以 $(a^{-1})^{-1} = a$, 关于两个元乘积的逆元,有 $(ab)^{-1} = b^{-1}a^{-1}$, 这是因为

$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}b = e.$$

定理 3 一个群的乘法适合消去律:a, b, b' 是群 G 的元,若 $ab = ab'$ 或 $ba = b'a$, 那么 $b = b'$

证 由 $ab = ab'$, 有

$$(a^{-1})ab = a^{-1}(ab'),$$

即

$$(a^{-1}a)b = (a^{-1}a)b',$$

所以

$$eb = eb',$$

$$b = b',$$

同理由 $ba = b'a$, 得到 $b = b'$. 定理得证.

定理 4 群定义中的(3), (4)两个条件,可以用如下条件来代替:对于群中任意元 a, b, 方程 $ax = b$, $ya = b$ 在这群中有解.

证 首先用群定义中的(1), (2)两条件与定理中的条件来推出群定义中的(3), (4)两个条件.

因为方程 $ya = a$ 有解,而 $ea = a$, 所以 e 是这一解,对任意元 b, $ax = b$ 有解 t,使得 $at = b$, $eb = e(at) = eat = at = b$, 所以

$$eb = b,$$

即左单位元存在，群定义中条件(3)成立.

又 $ya=e$ 有解，此解记为 a^{-1}，即 $a^{-1}a=e$，所以 a 的左逆元存在，群的定义中条件(4)成立.

反过由(3)，(4)两个条件推出定理中的条件：$ax=b$，$ya=b$ 在群中有解.

取 $x=a^{-1}b$，由(4)，a^{-1} 存在；由(1)，$a^{-1}b\in G$，G 的这个元显然是方程 $ax=b$ 的解，因为 $a(a^{-1}b)=(aa^{-1})b=eb=b$.

同样地，ba^{-1} 是方程 $ya=b$ 的解. 定理得证.

经常说 $a^{-1}b$ 是 a 左除 b 的商，ba^{-1} 是 a 右除 b 的商. 因为乘法与因子的顺序有关，所以 a 左除 b 与 a 右除 b，它们的商一般是不相等的. 于是群中，乘法这种运算是有逆运算的，叫做除法，并且对于除法来说，它也与乘法一样是闭合的.

由定理 4 的证明得知，群的定义也可以作如下描述：一个非空集合，假如它对于乘法和它们逆运算除法都是闭合的，并且乘法的结合律成立，那么它就成为群.

三、群表

一个乘集或一个群，假如其中任意两元的乘积已经知道，那么这乘集或这群的结合法也就完全知道，假如 $G=\{a_1, a_2, \cdots a_n, \cdots\}$. 可以用下面的表 4.1.1 来表示 G 中任意两元的结合法，叫做乘法表，当 G 是群时，这表叫做群表.

表 4.1.1

	a_1	$\cdots$	a_n	$\cdots$
a_1	a_1a_1	$\cdots$	a_1a_n	$\cdots$
$\vdots$	$\vdots$		$\vdots$	
a_n	a_na_1	$\cdots$	a_na_n	$\cdots$
$\vdots$	$\vdots$		$\vdots$	

表 4.1.2

e	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

例如，前面的例 4，$G=\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}-1&0\\0&-1\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix},\begin{pmatrix}-1&0\\0&1\end{pmatrix}\right\}$，

设 $a=\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$，$b=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$，$c=\begin{pmatrix}-1&0\\0&1\end{pmatrix}$，$e=\begin{pmatrix}1&0\\0&1\end{pmatrix}$，且 G 中元的结合法为矩阵的乘法，那么它的群表就是表 4.1.2：

从乘法表来看元素间的结合法非常明显，因此，给出一个群时，常常就给出它的群表，在群表的每行每列中，群的任意元必定出现一次而且只能出现一次. 交换群并且只有交换群才有关于主对角线对称的群表.

当群 G 为交换群时，常把结合法叫做加法，积 ab 记为 $a+b$，这时 G 叫做加群，单位元写成 0，叫做零元，a 的逆元写成 $-a$，叫做负 a.

第二节　变换群和置换群

一、变换群

集合 M 到 M 自己的映射，叫做 M 的一个变换，若这个映射是可逆映射，则这个映射叫做 M 的一一变换. 规定变换之间的一个运算叫做变换的乘法，也叫做变换的积，也就是说，任意两个变换 s，t 的乘积 st 是指先施行变换 t，再施行变换 s，设 a 是施行变换的对象，那么 $st(a)=s(t(a))$. 所以 st 仍然是一个变换.

把一个非空集合 M 的所有一一变换放在一起，作为一个集合

$$G=\{s,\ t,\ \tau,\ \lambda\cdots\},$$

那么对于变换的乘法，G 是否成为一个群呢？

定理 1　非空集合 M 的所有一一变换对变换的乘法构成一个群.

证　用 G 表示 M 的所有一一变换所构成的集合. 首先证明 s，t 是一一变换，st 也是一一变换.

设任意的 $a_1,\ a_2\in M,\ a_1\neq a_2$. 因为 t 是双射，则 $t(a_1)\neq t(a_2)$，从而 $s(t(a_1))\neq s(t(a_2))$，即 $st(a_1)\neq st(a_2)$；同时对任意 $b\in M$，由于 s 是双射，所以存在 $c\in M$ 使得 $s(c)=b$，再由 t 也是双射知存在 $a\in M$，使得 $t(a)=c$，于是 $s(t(a))=s(c)=b$，这就证明了 st 也是一一变换，所以 G 对变换乘法闭合.

又因为 $s,\ t,\ \tau\in G,\ a\in M,\ (st)\tau(a)=(st)(\tau(a))=s(t(\tau(a)))$；

$$s(t\tau)(a)=s(t\tau(a))=s(t(\tau(a))),$$

所以 $(st)\tau=s(t\tau)$，因此结合律成立. 恒等变换 I 是把每个施行对象仍然变成自己的映射，因此对任意的一一变换 s，都有 $Is=s$. 故 I 就是 G 的单位元. 对任意 $s\in G$，由于 s 是双射，故有一个 $s^{-1}\in G$，使得 $s^{-1}\cdot s=s\cdot s^{-1}=I$，所以 s 有逆元 s^{-1}. 综上所述，G 满足群定义中的 4 个条件，故 G 是一个群.

由以上定理可以定义变换群.

定义 1　设 M 是非空集合，由 M 的若干个一一变换对于变换的乘法构成的群叫做集合 M 的一个变换群.

例 1　设 A 为坐标平面的所有点构成的集合，令 $G=\{R_a\mid R_a$ 表示 A 中每个元素沿 y 轴方向平移 a 个单位，a 为任意实数$\}$，则 G 是集合 A 上的一个变换群，其中，当 $a\geqslant 0$ 时，R_a 表示 A 中每个元素沿 y 轴正向平移 a 个单位；当 $a<0$ 时，R_a 表示 A 中每个元素沿 y 轴负向平移 $|a|$ 个单位.

证 显然G的每个元素是集合A的一一变换，任取R_a，$R_b \in G$，则$R_aR_b = R_{a+b}$仍是A的一一变换，所以结合法闭合. 而且R_0为G的单位元；如果$R_a \in G$，则$R_{-a} \in G$使得$R_a \cdot R_{-a} = R_{-a} \cdot R_a = R_0$，所以$G$中每个元素有逆元. 至于$G$的乘法满足结合律是显然的，这是因为$(R_aR_b)R_c = R_{a+b} \cdot R_c = R_{a+b+c}$，$R_a(R_b \cdot R_c) = R_a(R_{b+c}) = R_{a+b+c}$，所以$(R_aR_b)R_c = R_a(R_bR_c)$. 这就证明了$G$是$A$的一个变换群.

例 2 设A是平面的所有点构成的集合，那么平面上绕一个定点的旋转可以看成A的一个一一变换，若G是包含所有绕一定点的旋转，那么G就是一个变换群.

证 用τ_θ表示转θ角的旋转，就有$\tau_{\theta_1} \cdot \tau_{\theta_2} = \tau_{\theta_1+\theta_2}$，结合法是闭合的，$(\tau_{\theta_1} \cdot \tau_{\theta_2})\tau_{\theta_3} = \tau_{\theta_1+\theta_2+\theta_3}$，$\tau_{\theta_1}(\tau_{\theta_2} \cdot \tau_{\theta_3}) = \tau_{\theta_1+\theta_2+\theta_3}$，所以结合律成立. G的单位元为τ_0，$\tau_\theta^{-1} = \tau_{-\theta}$为$\tau_\theta$的逆元. 所以$G$是一个群.

变换群一般不是交换群.

例如s是平面的一个平移，它把原点$(0, 0)$平移到$(1, 0)$，t是绕原点转$\frac{\pi}{2}$的旋转，那么s，t都是上例的集合A的一一变换，但是

$$st:(0, 0) \to (1, 0)$$

$$ts:(0, 0) \to (0, 1)$$

$$st \neq ts$$

故非交换群是存在的.

二、置换群

变换群的一种特例，叫做置换群，在群里占一个很重要的地位. 这种群还有一个特点，就是它们的元可以用一种很具体的符号来表示，使得在这群里的计算比较简单，下面对置换群进行讨论.

定义 2 **如果M是一个有限集合，M的一个一一变换叫做M的一个置换. 假如M是元数为n的有限集，M的所有置换所构成的群叫做n次对称群，用S_n表示，由若干个置换构成的群叫做一个置换群.**

由初等代数的知识，n个元的置换一共有$n!$个. 因为要作n个元的一个置换，就要替每一个元选定一个对象，例如替a_1选定对象时，有n个可能，选定之后，a_2选时有$n-1$个可能，这样下去就可以得到$n(n-1)\cdots 2 \cdot 1 = n!$个不同的置换. 因此有：

定理 2 **n次对称群S_n的阶是$n!$.**

设$M = \{a_1, a_2, \cdots, a_n\}$，对任意$f \in S_n$，用一张表来表现$M$的每个元素

$a_i(1 \leqslant i \leqslant n)$ 在 f 之下的像$f(a_i)$，也即可将 f 表示为

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ f(a_1) & f(a_2) & \cdots & f(a_n) \end{pmatrix},$$

当然，对于每个置换 f，也可以表示为

$$\begin{pmatrix} a_2 & a_1 & \cdots & a_n \\ f(a_2) & f(a_1) & \cdots & f(a_n) \end{pmatrix}.$$

所以，在表示 f 时，M 中的元素只起到记号的作用，为了更简洁地表示 S_n 的元素，不妨假定 $M=\{1, 2, \cdots, n\}$，此时 S_n 的元素 f 就可以记为

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ f(1) & f(2) & \cdots & f(n) \end{pmatrix}.$$

所以，有时 M 的置换叫做排列. 例如 $s=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$ 为$\{1, 2, 3, 4\}$的一个排列.

例 3 写出 S_3 的所有元素，并证明 S_3 是非交换群.

解 S_3 共有 6 个元：

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

又取

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix};$$

则

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

故 S_3 不是交换群.

定义 3 **设 $f \in S_n$，如果 $f(a_{i_1})=a_{i_2}$，$f(a_{i_2})=a_{i_3}$，$\cdots$，$f(a_{i_{k-1}})=a_{i_k}$，$f(a_{i_k})=a_{i_1}$，而 M 中其余的元素都不变动，这样的一个 f 就叫做一个k 阶循环置换(循环排列)，用 $(a_{i_1}a_{i_2}\cdots a_{i_k})$ 或 $(a_{i_2}a_{i_3}\cdots a_{i_k}a_{i_1})$ 表示. 若 $f=(a_{i_1}a_{i_2})$，则 f 叫做一个对换.**

显然，一阶循环置换就是 S_n 的恒等置换 I，所以 $I=(1)$.

例 4 $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}=(1\ 2\ 3)=(2\ 3\ 1)=(3\ 1\ 2)$,

$$\begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}=(1)=(2)=(3)=(4)=(5),$$

$$\begin{pmatrix}1&2&3&4&5\\2&1&3&4&5\end{pmatrix}=(1\ 2),\text{这是一个对换}.$$

$$\begin{pmatrix}1&2&3&4&5\\3&5&4&1&2\end{pmatrix}=\begin{pmatrix}1&2&3&4&5\\3&2&4&1&5\end{pmatrix}\begin{pmatrix}1&2&3&4&5\\1&5&3&4&2\end{pmatrix}$$
$$=(1\ 3\ 4)(2\ 5).$$

一般有：

定理 3　每一个 n 元的置换都可以表示成若干个互相没有公共数字的循环置换的乘积.

证　用归纳法证明. 设 $f\in S_n$，如果 $f=I$，则 $f=(1)$，所以此时定理成立. 假定当 f 的变动元素的个数小于等于 $r-1$ 时定理成立，下面证明 f 的变动元素个数为 r 时定理也成立.

任取置换 f 的一个变动元素 i_1，假定 $f(i_1)=i_2$，$f(i_2)=i_3$，…，$f(i_k)=i_{k+1}$，…. 因 A 为有限集，所以必有 i_1，…，i_k 互不相同，而 i_{k+1} 与 i_1，…i_k 中某一个相同. 现在要证 $i_{k+1}=i_1$. 因为

$$f(i_1)=i_2,\ f(i_2)=i_3,\ \cdots,\ f(i_{k-1})=i_k,$$

所以 i_2，i_3，…，i_k 分别是 i_1，i_2，…，i_{k-1} 在 f 之下的像，由于 f 是一个置换，故 i_k 在 f 之下的像不能是 i_2，i_3，…，i_k 中任何一个，这就是说 i_k 在 f 之下的像 i_{k+1} 只能是 i_1.

又因为 f 变动 r 个元素，所以 $k\leqslant r$，如果 $k=r$，则 f 就是一个循环置换. 如果 $k<r$，那么

$$f=\begin{pmatrix}i_1&i_2&\cdots&i_{k-1}&i_k&i_{k+1}&\cdots&i_r&i_{r+1}&\cdots&i_n\\i_2&i_3&\cdots&i_k&i_1&i_{k+1}^{(1)}&\cdots&i_r^{(1)}&i_{r+1}&\cdots&i_n\end{pmatrix}$$

$$=(i_1,i_2,\cdots,i_k)\begin{pmatrix}i_1&\cdots&i_k&i_{k+1}&\cdots&i_r&i_{r+1}&\cdots&i_n\\i_1&\cdots&i_k&i_{k+1}^{(1)}&\cdots&i_r^{(1)}&i_{r+1}&\cdots&i_n\end{pmatrix}$$

$$=(i_1i_2\cdots i_k)g,$$

其中
$$g=\begin{pmatrix}i_1&\cdots&i_k&i_{k+1}&\cdots&i_r&i_{r+1}&\cdots&i_n\\i_1&\cdots&i_k&i_{k+1}^{(1)}&\cdots&i_r^{(1)}&i_{r+1}&\cdots&i_n\end{pmatrix}.$$

由于 g 只使 $r-k(r-k<r)$ 个元变动，由归纳法的假定，g 可以写成没有公共数字的循环置换的乘积

$$g=\eta_1\eta_2\cdots\eta_m.$$

在这 η_i 之中，$i_1, i_2, \cdots, i_k$ 不会出现，否则

$$\eta_i = (\cdots i_p \quad i_q \cdots), \qquad p \leqslant k.$$

那么 i_p 同 i_q 不会再在其余的 η 中出现，g 也必然会使 $i_p \to i_q$，但 g 是使 i_p 不变的，这是一个矛盾. 这样 f 就是没有公共数字的循环置换的乘积

$$f = (i_1 i_2 \cdots i_k) \eta_1 \eta_2 \cdots \eta_m.$$

推论　每个置换都可以表示为若干个对换的乘积.

证　先证明每个循环置换可表示为若干个对换的乘积. 设 f 是一个 k 阶循环置换，$f = (i_1 i_2 \cdots i_k)$.

当 $k = 1$ 时，$f = (i_1) = (i_1 i_2)(i_2 i_1)$；

当 $k \geqslant 2$ 时，$f = (i_1 i_2 \cdots i_k) = (i_1 i_k)(i_1 i_{k-1}) \cdots (i_1 i_2)$.

而由定理 3，每个置换又可以表示为若干个没有公共数字的循环置换的乘积，故推论得证.

例 4　用循环置换的方法写出 S_4 的全体元素.

解　S_4 有 24 个元素，它们是：

(1)；

(12)，(13)，(14)，(23)，(24)，(34)；

(123)，(124)，(132)，(142)，(134)，(143)，(234)，(243)；

(1234)，(1243)，(1324)，(1342)，(1423)，(1432)；

(12)(34)，(13)(24)，(14)(23).

第三节　子　　群

一、子群的定义及其判别法

在研究一个群的构造时，通常通过群的子集去刻画整个群的性质，为此需要介绍子群的概念，子群的概念本身也揭示了群的结构性质.

定义 1　**一个群 G 的非空子集 H，如果 H 对于 G 的结合法成为一个群，那么 H 就叫做 G 的一个子群，记为 $H \leqslant G$. 如果 $H \neq G$，则 H 是 G 的真子群，记为 $H < G$.**

例 1　整数集对于加法形成的群是有理数集对于加法形成的群的子群.

例 2　设 $H = \{(1), (123), (132)\}$，则 H 是 $S_3 = \{(1), (12), (13), (23), (123)(132)\}$ 的一个子群.

例 3　G 是一个群，e 为 G 的单位元，那么 $H = \{e\}$ 和 G 是 G 的子群，一般这两个子群叫做 G 的平凡子群，G 的其他子群叫做 G 的真子群.

一个群的子群有以下性质：

定理 1　设 H 是群 G 的子群，则

(1) H 的单位元就是 G 的单位元；

(2) H 中元 a 的逆元也就是 a 在 G 中的逆元.

证　(1) 设 e_1 及 e 分别是 H 和 G 的单位元，则

$$ee_1 = e_1e_1 = e_1.$$

由于方程 $ye_1 = e_1$ 在 G 中有且只有一个解，故 $e = e_1$.

(2) 设 $a \in H$，a 在 H 中的逆元是 a'，a 在 G 中的逆元为 a^{-1}，则

$$aa' = aa^{-1} = e.$$

由 G 的乘法消去律知 $a' = a^{-1}$.

下面讨论子集成为子群的条件.

定理 2　群 G 的非空子集 H 成为其子群的充分必要条件是：

(1) 若 $a, b \in H$，则 $ab \in H$；

(2) 若 $a \in H$，则 $a^{-1} \in H$.

证　必要性是很显然的，因为 H 成群，(1)(2)是必然成立的，下面来证明充分性. 由于(1)，H 是闭合的，结合律在 G 中成立，当然在 H 中成立. 由(2)，假如 H 包含 a，H 也包含 a^{-1}，因此 H 包含 $a \cdot a^{-1} = e$，即 $e \in H$，于是 H 成群，所以定理成立.

把上面定理的两个条件合并成一个，就得到：

定理 3　群 G 的非空子集 H 成为其子群的充分必要条件是：若 $a, b \in H$，那么 $ab^{-1} \in H$.

证　先证明由定理 2 中(1)，(2)成立，得到定理 3 成立. 假定 $a \in H$，$b \in H$，由(2)得到 $b^{-1} \in H$，再由(1)得到 $ab^{-1} \in H$. 反过来，如果 $a \in H$，$b \in H$，则 $a \cdot a^{-1} = e \in H$，也就有 $e \cdot a^{-1} \in H$，即 $a^{-1} \in H$. 同理 $b^{-1} \in H$，所以 $a \cdot (b^{-1})^{-1} = ab \in H$.

特别地，当 H 是非空有限子集时，它成为子群的充分必要条件是定理 2 中的条件(1)：$a, b \in H$，$a \cdot b \in H$. 因为这时，可以用消去律来代替群定义中条件(3)，(4)，消去律在 G 中成立，因此在 H 中也同样成立.

例 4　由实数组成的所有 n 阶满秩矩阵对乘法成为群，叫做 $\mathbf{R}$ 上的 n 阶线性群，用 $GL(n, \mathbf{R})$ 表示，这里的 $\mathbf{R}$ 表示实数集. 那么 $GL(n, \mathbf{R})$ 中所有行列式为 1 的矩阵，成为其子群，用 $SL(n, \mathbf{R})$ 表示. 这是因为 $A, B \in SL(n, \mathbf{R})$. $|A| = 1$，$|B| = 1$，$|AB| = |A||B| = 1$，故 $AB \in SL(n, \mathbf{R})$，$|A^{-1}| = 1$，故 $A^{-1} \in SL(n, \mathbf{R})$. 所以 $SL(n, \mathbf{R})$ 是 $GL(n, \mathbf{R})$ 的子群.

关于子群,我们还有：

定理 4 **设 G 是一个群，$H_i(i\in I)$ 为 G 的子群，则 $H=\bigcap\limits_{i\in I}H_i$ 是 G 的子群.**

证 设 $a,\ b\in H$，则 $a,\ b$ 属于每个 H_i，所以 ab^{-1} 属于每个 H_i，即 $ab^{-1}\in\bigcap\limits_{i\in I}H_i$，由定理 3 知 H 也是 G 的子群.

定理 4 说明一个群的任意多个子群的交集仍然是一个子群，但需要注意的是任意两个子群的并集一般不成群.

二、交代群

下面介绍对称群的重要子群——交代群.

在本章第二节中讨论置换时知道，每个置换可以表示为若干个对换的乘积，一般说来，其表示的方法不是惟一的.

例如 $(123)=(13)(12)=(32)(12)=(12)(31)$. 虽然如此，却有下面的定理.

定理 5 **如果把一个置换用对换的乘积来表示，那么对换因子的个数是偶数或奇数是一定的.**

证 设 n 个数 $a_1,\ a_2,\ \cdots,\ a_n$ 的函数

$$F=\begin{vmatrix}1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\\ a_1^2 & a_2{}^2 & \cdots & a_n^2\\ \vdots & \vdots & & \vdots\\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1}\end{vmatrix}=\prod_{1\leqslant i<j\leqslant n}(a_i-a_j)$$

$$\begin{aligned}&=(a_{n-1}-a_n)\\&\cdot(a_{n-2}-a_n)(a_{n-2}-a_{n-1})\\&\cdot\cdots\\&\cdot(a_1-a_n)(a_1-a_{n-1})\cdots(a_1-a_2)\end{aligned}$$

F 也可写成

$$F=(a_k-a_l)\prod_{i\neq k,l}(a_i-a_k)(a_i-a_l)f,$$

这里 f 不包括 a_k 及 a_l. 在 F 上施行对换 (a_ka_l)，这时 f 及 $(a_i-a_k)(a_i-a_l)$ 都不变动，但 a_k-a_l 变了符号，因此 F 换成了 $-F$. 也就是说，在 F 上施行任一个对换，F 就改变了符号. 因此在 F 上施行偶数个对换，结果仍是 F，在 F 上施行奇数个对换，结果是 $-F$. 而在 F 上施行一个置换的结果是一定的，因此一个置换

不能同时表为偶数个对换的乘积，又表为奇数个对换的乘积，因此定理得证.

定义 2 一个置换，它的对换因子的个数如果是偶数，就叫做偶置换，如果是奇数就叫做奇置换.

一个偶置换与一个奇置换的乘积是奇置换，两个偶置换或两个奇置换的乘积都是偶置换. 恒等置换是偶置换. 一个循环置换，它所含数字的个数如果是偶数就是奇置换，如果是奇数就是偶置换，这是因为

$$(12345\cdots n) = (12)(23)\cdots(n-1\quad n).$$

两个偶置换的乘积仍是偶置换，而且恒等置换是偶置换，因此有

定义 3 n **次对称群** S_n **中所有偶置换成为** S_n **一个子群，叫做** n **次交代群，用** A_n **表示.**

A_n 是 S_n 中一个非常重要的子群. 当用 S_n 中某一奇置换乘其中所有偶置换时就得到互异的奇置换，因此 S_n 中偶置换不多于奇置换. 同样地，S_n 中奇置换也不会多于偶置换，所以 S_n 中偶置换与奇置换个数相等. 又因为 $|S_n| = n!$，所以有

$$|A_n| = \frac{n!}{2}.$$

例 5 $S_3 = \{(1), (12), (13), (23), (123), (132)\}$,

$A_3 = \{(1), (123), (132)\}$,

$A_4 = \{(1), (12)(34), (13)(24), (14)(23), (123), (124), (134), (234), (132), (142), (143), (243)\}$.

三、子集生成的子群

一个群到底有多少个子群？子群的构造如何？这是群论中主要问题之一. 下面学习一种找一个子群的一般方法.

设 M 为群 G 的一个非空子集，当然 M 不一定成群，因为 M 中任意两元的积及任意元的逆虽然都在 G 中，但却不一定都在 M 中. 但是可以把 M 扩大一些，而得到一个包含 M 的子群.

把 M 的元以及这些元的逆作各种乘积，例如 ab, $a^{-1}c$, $b^2a^{-1}c^{-1}$, d 等等，作成一个集合 H，让它刚好包含这些乘积. 因为两个这样的乘积仍是这样的乘积，一个这样乘积的逆元也是这样的乘积，故 H 就是一个子群. 换句话说，H 就是 M 添加其元的乘积及逆，当添加一旦成群后，就不再添加. 所以 H 就是 G 中包含 M 的最小子群.

定义 4 如上得到的 H **叫做由** M **生成的子群，用** $\langle M\rangle$ **表示它.**

例 6 由(1234)，(12)(34)生成的群

$$D_8 = \langle (1234), (12)(34) \rangle$$
$$= \{(1), (13), (24), (12)(34), (13)(24), (14)(23), (1432), (1234)\}.$$

D_8 是一个 8 元群，叫做二面体群，是一个常见的群.

第四节 循 环 群

一、循环群的概念及构造

第三节介绍了由子集生成的子群，假如取一个只包含一个元 a 的子集，那么由它生成的子群又会有什么特性呢？这就是这一节要介绍的循环群.

定义 1 设 G 是一个群，如果存在 G 的一个元 a 使得 $G = \langle a \rangle$，则 G 叫做一个循环群，并且 a 叫做 G 的生成元.

也常用(a)表示$\langle a \rangle$，此时 $G = (a)$ 中元的一般构造是 a 的所有乘幂 a^n，而且这时 $a^m \cdot a^n = a^{m+n}$， $a^0 = e$， $a^{-n} = (a^{-1})^n$.

例 1 整数加群是一个循环群，1，-1 都是它的生成元，即

$$\mathbf{Z} = (1) = (-1).$$

例 2 A_3 是元数为 3 的循环群，(123)，(132)都是它的生成元，即

$$A_3 = ((123)) = ((132)).$$

循环群在群中构造是最简单的，并且是最基本的. 下面讨论循环群(a)的构造.

假定 a 的所有乘幂都互不相等，即当 $i \neq j$ 时，$a^j \neq a^i$，这时循环群(a)是由下面的元组成的：

$$\cdots, a^{-3}, a^{-2}, a^{-1}, a^0, a^1, a^2, \cdots$$

它是无限群.

假定 a 的乘幂中有相等的，并且 $h > k$ 时，$a^h = a^k$，这时有

$$a^{h-k} = e, \qquad h-k > 0,$$

这就是说 a 的正的次幂中有一个是单位元. 如果 n 是使 $a^n = e$ 成立的最小正整数，那么$a^0, a^1, \cdots, a^{n-1}$彼此互不相同. 这是因为假如存在 i, j, $a^i = a^j$，那么有 $a^{i-j} = e$, $0 < i-j < n$. 这就与前面假设的 n 是最小正整数相矛盾. 再假如把任意正整数 m 写成

$$m = qn + r \qquad (0 \leqslant r < n),$$

那么就有　　$a^m = a^{qn+r} = a^{qn} \cdot a^r = (a^n)^q \cdot a^r = e \cdot a^r = a^r$

所以 a 的任意乘幂都与 a^0, a^1, a^2, $\cdots$, a^{n-1} 中某一个相等. 因此循环群(a)的元只有 n 个：

$$a^0, a^1, a^2, \cdots, a^{n-1}.$$

它是一个有限群，元数是 n.

定义 2　假如 a 的任意两个乘幂都不相等，就说 a 的阶是无穷的. 假如 a 的乘幂中有相等的，n 是使 $a^n = e$ 成立的最小正整数，就说 a 的阶数是 n.

例 1 中，1 及 -1 的阶都是无穷的，例 2 中 A_3 的元(123)的阶数是 3，这是因为 $(123)^3 = (1) = e$.

群的单位元的阶数是 1.

关于元的阶数还有以下重要结论：

若 a 的阶是无穷的，那么当 $a^n = e$ 时，$n = 0$. 若 $a^r = a^s$ 成立，其充分必要条件是 $r = s$.

若 a 的阶数是 n, $a^n = e$, 若 $a^m = e$, 则 m 应是 n 的倍数，即$n|m$. 这是因为 $m = qn + r$, $0 \leqslant r < n$. 于是

$$r = m - qn, \quad a^r = a^{m-qn} = a^m \cdot (a^n)^{-q} = e \cdot e = e,$$

所以 $a^r = e$. 又因为 a 的阶数是 n，所以 $r = 0$，故 $m = qn$. 而 $a^r = a^s$ 的充分必要条件是 $r \equiv s(n)$. 这是因为 $a^r = a^s$，所以 $r - s = qn$，故 $r \equiv s(n)$. 反过来，当 $r \equiv s(n)$ 时，$r = qn + s$，当然有 $a^r = a^s$.

于是由以上的讨论，有

定理 1　一个循环群(a)，假如 a 的阶是无穷的，那么(a)是无限群：

$$(a) = \{\cdots, a^{-2}, a^{-1}, a^0, a^1 \cdot a^2 \cdots\}.$$

假如 a 的阶数是 n，那么它是元数等于 n 的有限群：

$$(a) = \{a^0, a^1, a^2, \cdots, a^{n-1}\}, \quad a^n = a^0.$$

通过以上的讨论，就清楚了循环群的基本构造，下面讨论循环群的子群.

二、循环群的子群

定理 2　循环群(a)的子群是循环群.

证　设 G 是由元 a 生成的循环群，$G = (a)$，H 是 G 的子群. 如果 $H = \{e\}$，即 H 是单位元群，那么显然 H 是一个循环群.

如果 $H \neq \{e\}$，任取 $a^k \in H$，且 $a^k \neq e$，由于 H 是G 的子群，故 a^k 的逆元 $a^{-k} \in H$，由 $a^k \neq e$ 知 $k \neq 0$，所以$-k$与 k 之中必有一个为正整数，故 H 中必

含有幂 $k>0$ 的元 a^k. 假设 a^m 是 H 中 a 的最小正幂,下面证明 $H=(a^m)$.

设 a^s 为 H 中任意元,由 $s=tm+l$, $0\leqslant l<m$, 有 $a^l=a^{s-tm}=a^s\cdot(a^m)^{-t}$ 也是 H 中的元,这与 m 的取法矛盾,所以只有 $l=0$, 即 $a^s=(a^m)^t$, 这就是说 H 中任意元是 a^m 的乘幂,也即 H 只含 a^m 的任意乘幂. 故 H 是由 a^m 生成的循环群,即 $H=(a^m)$. 定理得证.

因为 G 中任意元的 m 乘幂都是 a^m 的乘幂,所以 H 又可以看成是由 G 中各元的 m 乘幂组成的子群.

假如 a 的阶是无穷的, a^m 的阶也是无穷的,于是 H 由下面的元组成:

$$(a^m)^0=e,\ a^{\pm m},\ a^{\pm 2m},\ a^{\pm 3m},\ \cdots$$

这时它是无限群.

假如 a 的阶数是 n, 即 $a^n=e$, 因为 a^m 是 H 中 a 的最小正幂,而 $a^n\in H$, 所以 $m|n$, 令 $n=qm$, 因为 $(a^m)^q=a^{mq}=a^n=e$, 所以 a^m 的阶为 q, 于是 H 只包含下面的 q 个元:

$$a^0,\ a^m,\ a^{2m},\ \cdots,\ a^{(q-1)m},$$

所以这时 H 是元数为 q 的有限群. G 中元数为 q 的子群显然只是由 a^m 生成的,因此 G 只有惟一一个 q 元子群.

于是,无限循环群有无穷多个子群. n 元循环群的子群的个数等于 n 中互异正因数的个数,这是因为 G 的子群 H 的元数为 $\frac{n}{m}$, a^m 为 H 中 a 的最小正幂.

循环群的两个子群是否相等,有下面的定理:

定理 3　循环群 $G=(a)$ 是无限群,那么它的两个子群 (a^r), (a^s) 相等的充分必要条件是 $r=\pm s$.

证　若 $r=\pm s$, 由于 G 是无限群, $(a^r)=(a^s)$ 是成立的,因而充分性是成立的.

对于必要性,若 $(a^r)=(a^s)$, 那么 (a^r) 中的元 a^r 应由 a^s 生成,即 $a^r=(a^s)^k=a^{sk}$, 故 $r=sk$, 同理 $s=rl$. 于是 $s=skl$, 得到 $kl=1$, 由于这里的 k, l 均是整数,故 $k=l=\pm 1$, 于是 $r=\pm s$. 定理得证.

定理 4　假如 $G=(a)$ 是有限的 n 阶循环群,那么它的两个子群 (a^r), (a^s) 相等的充分必要条件是 $(r,n)=(s,n)$.

证　对于充分性,因为 $(a^r)=(a^s)$, 所以 $a^r=(a^s)^k$, 于是 $a^r\cdot a^{-sk}=e$, 所以 $a^{r-sk}=e$. 又因为 (a) 的阶为 n, 所以 $r-sk=nh$, 于是 $(s,n)|r$, 因此 $(s,n)|(r,n)$, 同理有 $(r,n)|(s,n)$, 于是 $(s,n)=(r,n)$.

对于必要性,若 $(r,n)=(s,n)$, 那么 $(s,n)|r$, 于是 $r-sk=nh$ 中的 k, h 这两个整数是存在的,即 $r=sk+nh$ 成立,于是

$$a^r = (a^{sk+nh}) = a^{sk} \cdot a^{nh} = (a^s)^k,$$

所以 $a^r \in (a^s)$，同理有 $a^s \in (a^r)$，故 $(a^r) = (a^s)$. 定理得证.

第五节　子群的陪集

一、子集的乘积

这一节里，要利用群 G 的一个子群来作 G 的分类，先介绍群的两个子集乘积的概念.

定义 1　设 A, B 是群 G 的两个非空子集，集合 $\{ab \mid a \in A, b \in B\}$ 叫做 A 与 B 的乘积，或简称为 A, B 的积，记作 AB，即

$$AB = \{ab \mid a \in A, b \in B\}.$$

例如，取群 S_3 的子集 $A = \{(12)\}$, $B = \{(12), (13)\}$，则

$$AB = \{(12)\}\{(12), (13)\} = \{(1), (132)\}.$$

对于群的三个子集 A, B, C 的乘积，由定义有 $A(BC) = (AB)C$，这是因为当任意 $a \in A$, $b \in B$, $c \in C$ 时，$a(bc) \in A(BC)$，而 $a(bc) = (ab)c \in (AB)C$，所以 $A(BC) \subseteq (AB)C$，同理 $(AB)C \subseteq A(BC)$，所以 $A(BC) = (AB)C$. 这也就是说，群中子集的乘积是满足结合律的. 当 H 是子群时，$HH = H$.

当 A, B 是群 G 的子群时，AB 不一定成群，那么 AB 成群的条件是什么呢？有下面的定理.

定理 1　群 G 的子群 A, B 的乘积 AB 成群的充分必要条件是 A, B 可以交换，即 $AB = BA$.

证　假设 $AB = BA$，AB 中任意元 ab 的逆元 $b^{-1}a^{-1}$ 就在 BA 中，由于 $AB = BA$，故 ab 的逆在 AB 之中. 又因为 $ABAB = AABB = AB$，这表明 AB 中任意两元的乘积仍在 AB 中，于是 AB 成群.

当 AB 成群时，设 a, b 分别为 A, B 之中的任意元. 由于 BA 中元 ba 是 AB 中的元 $a^{-1}b^{-1}$ 的逆，故 $ba \in AB$，所以 $BA \subseteq AB$. 又因为 $(ab)^{-1} \in AB$，令 $(ab)^{-1} = a_1b_1$，于是 $ab = b_1^{-1}a_1^{-1} \in BA$，故 $AB \subseteq BA$，所以 $AB = BA$，即 A, B 可交换. 定理得证.

以上定理需要说明的是，$AB = BA$，实际上是它们互相包含，因为 a, b 是 A, B 中元时，ab 不一定等于 ba，而是 $ab = b'a'$, $a' \in A$, $b' \in B$，或者 $ab = b''a''$, $a'' \in A$, $b'' \in B$.

当 G 是交换群时，$AB = BA$ 是一定成立的，故交换群的两个子群的乘积仍

是 G 的一个子群.

特别地,当 B 或 A 是包含一个元的子集时,将 AB 记为

$$AB = aB, \quad A = \{a\}.$$

二、左(右)陪集及陪集分解

定义 2　设 H 是群 G 的一个子群,a 是 G 的任意一个元,那么集合 $aH(Ha)$,叫做 G 中 H 的左(右)陪集.

当 G 为加群时,G 中 H 的左(右)陪集记为 $a+H(H+a)$.

例 1　$G = S_3$,取 G 的子群 $H = \{(1), (12)\}$,

H 的左陪集

$(13)H = \{(13), (123)\}$,　$(23)H = \{(23), (132)\}$, $(1)H = \{(1), (12)\}$;

H 的右陪集

$H(13) = \{(13), (132)\}$,　　$H(23) = \{(23), (123)\} = H(132)$.

一般说来,群的子群的左、右陪集不一定一致,但当 G 是交换群时,一定有 $aH = Ha$,故在交换群里,陪集就不用分左右了.

对于群 G 的子群的不同陪集 aH 与 bH 有:

定理 2　设 H 是群 G 的一个子群,任取 $a, b \in G$,则下列条件等价:

(1) $a^{-1}b \in H$;

(2) $b = ah$, $h \in H$;

(3) $b \in aH$;

(4) $aH = bH$.

证　(1)⇒(2):假设(1)成立,即 $a^{-1}b \in H$,则存在 $h \in H$,使得 $a^{-1}b = h$,所以 $b = ah$ 成立.

(2)⇒(3):假设(2)成立,即存在 $h \in H$,使得 $b = ah$,由 aH 的定义知,$b \in aH$,即(3)成立.

(3)⇒(4):假设(3)成立,即 $b \in aH$,则存在 $h \in H$,使得 $b = ah$,任取 $c \in bH$,则有 $c = bh_1$, $h_1 \in H$,所以 $c = ahh_1 = a(hh_1) \in aH$,所以 $bH \subseteq aH$,同理 $aH \subseteq bH$,因此 $aH = bH$,于是(4)成立.

(4)⇒(1):假设(4)成立,即 $aH = bH$,显然会有 $b = be \in bH$,所以 $b \in aH$,因而存在 $h \in H$,使得 $b = ah$,故 $a^{-1}b = h \in H$,即 $a^{-1}b \in H$.

综上所述,(1),(2),(3),(4)等价.

再分析例 1,写出 S_3 的子群 H 的全体右陪集为

$$H(1) = \{(1), (12)\} = H(12),$$

$$H(13) = \{(13), (132)\} = H(132),$$

$$H(23) = \{(23), (123)\} = H(123).$$

如此这样，子群 H 把整个群 G 分成 $H(1)$，$H(13)$，$H(23)$ 三个不同的右陪集，这三个右陪集合在一起显然正是 G，因此可以说它们是 G 的一个分类. 对于左陪集也有类似情形.

对于一般的情况，因为 G 中任意元 a 一定会出现在 H 的左陪集 aH 中，假如左陪集 aH，bH 有公共元 $ah_1 = bh_2$，那么 $a^{-1}b = h_1h_2^{-1} \in H$，由定理 2 有 $aH = bH$. 这就是说，G 中任意两个左陪集或者重合或者没有公共元. 于是 G 可以分解为若干个互异的左陪集 a_iH，$i = 1, 2, \cdots$，即

$$G = a_1H \cup a_2H \cup \cdots \cup a_nH \cup \cdots.$$

以上的分解除了 a_iH 的顺序外是惟一的，但 a_i 不是惟一的.

同理，G 也可以惟一地分解为若干互异的右陪集 Hb_i，$i = 1, 2, \cdots$，即

$$G = Hb_1 \cup Hb_2 \cup \cdots.$$

一般地，用同一个子群把群分解为左陪集与右陪集时，其结果是不相同的，关于群的左、右陪集分解有以下重要关系.

定理 3 **设 H 是群 G 的子群，$G = a_1H \cup a_2H \cup \cdots \cup a_nH \cup \cdots$，则**

$$G = Ha_1^{-1} \cup Ha_2^{-1} \cup \cdots \cup Ha_n^{-1} \cup \cdots.$$

证 要证明 G 可以进行 $G = Ha_1^{-1} \cup Ha_2^{-1} \cup \cdots$ 的右陪集分解，则只要证明任意元 $g \in G$，而 g 在某个右陪集 Ha_i^{-1} 中，并且任意两个右陪集 Ha_i^{-1}，Ha_j^{-1} 都不相等.

首先因为 $g \in G$，故 $g^{-1} \in G$，于是 g^{-1} 在 H 的某个左陪集 a_iH 中，于是存在 $h \in H$，使得 $g^{-1} = a_ih$，所以 $g = h^{-1}a_i^{-1} \in Ha_i^{-1}$.

如果 H 的两个右陪集 Ha_i^{-1} 与 Ha_j^{-1} 相等，由定理 2，$a_i^{-1}a_j \in H$，因而有 $a_iH = a_jH$，而这是与假设不合的，所以 Ha_i^{-1} 与 Ha_j^{-1} 是不相等的右陪集. 因此定理得证.

三、Lagrange(拉格朗日)定理

由定理 3，我们得知 $\{a_1, a_2, \cdots\}$ 与 $\{a_1^{-1}, a_2^{-1}, \cdots\}$ 的元有相等的数目，故一个子群 H 的左陪集的个数与右陪集个数是相等的，或者无限大，或者有限且相等.

定义 3 **一个群 G 的一个子群 H 的相异左(右)陪集的个数，叫做 H 在 G 中的指标，用记号 $|G:H|$ 表示.**

例 2 交代群 A_4 对于 Klein(克莱茵)四元群 $B_4 = \{(1), (12)(34),$

$(13)(24), (14)(23)\}$ 的左陪集分解为 $A_4 = B_4 \cup (234)B_4 \cup (243)B_4$；

右陪集分解为 $A_4 = B_4 \cup B_4(234) \cup B_4(243)$. 而且 $|A_4 : B_4| = 3$.

例 3 G 是整数加群，取子群 $H = \{2n \mid n \in \mathbf{Z}\}$，$G$ 的 H 的右陪集分解为 $G = (H+0) \cup (H+1)$，$|G : H| = 2$.

定理 4 （Lagrange 定理）**设 H 是有限群 G 的一个子群，则 H 的阶整除 G 的阶.**

证 因为 G 可以分解成 $|G : H|$ 个 H 的左陪集，而每个左陪集所含元素的个数与 H 的个数是相同的，即每个左陪集有 $|H|$ 个元，因此有

$$|G| = |H||G : H|,$$

那么 $|H|$ 当然整除 $|G|$. 定理得证.

由定理 4，可以得到以下重要推论.

推论 1 **有限群 G 的每一个元的阶数整除 G 的阶.**

证 设 a 为 G 的元，由 a 生成一个阶为 n 的 G 的循环子群 H，由定理 4，有 $|H|$ 整除 $|G|$，即 n 整除 $|G|$.

推论 2 **设 G 是一个有限群，G 的阶是一个素数 p，则 G 没有异于单位元群的真子群.**

证 假定 H 是 G 的一个子群，则 $|H|$ 整除 $|G|$，已知 $|G|$ 是一个素数 p，所以 $|G|$ 的因数只有 1 和 p，因而 $|H|$ 为 1 或 p，从而推论得证.

推论 3 **一个素数阶的有限群一定是循环群.**

证 任取 $a \in G$，则 $|(a)|$ 整除 $|G|$ 得 p，p 为素数，又 $|(a)| > 1$，故 $|(a)| = p$，所以 $G = (a)$.

p^n 元循环群只有 n 个真子群，一般元数是素数 p 的幂的群叫做p-群，例如 B_4 是一个 2-群.

Lagrange 定理提供了一个确定有限群的子群的方法.

例 4 求出 S_3 的所有子群.

因为 $|S_3| = 6$，由 Lagrange 定理，S_3 的子群的阶只能是 1, 2, 3, 6. S_3 的 1 阶子群是 $H_1 = \{(1)\}$；S_3 的 2 阶子群有：$H_2 = \{(1), (2)\}$，$H_3 = \{(1), (13)\}$，$H_4 = \{(1), (23)\}$；S_3 的 3 阶子群有：$H_5 = \{(1), (123), (132)\}$；$S_3$ 的 6 阶子群是 S_3 本身. 因此 S_3 有 6 个子群.

但是，Lagrange 定理的逆不成立，即并非对 $|G|$ 的每一个正因数 m 都存在有限群 G 的 m 阶子群. 例如 A_4 是 12 阶群，6 是 12 的因数，但是 A_4 不含 6 阶子群. Lagrange 定理的逆对于循环群是成立的，而且对于交换群也是成立的. 关于以上的论述，下面有定理：

定理 5 （Sylow（西洛）第一定理）**假定有限群 G 的元数是 n，p 是素数，并**

且$p^r \mid n$，那么G有p^r元子群，当r为最大幂指数时，这p^r元子群叫做G属于p的Sylow子群，简称为p-Sylow子群.

定理5的证明留在以后的章节中给出.

例如A_4的元数$n = 12 = 2^2 \cdot 3$，那么它的2-Sylow子群是B_4，它的3-Sylow子群是由(123)，(134)，(142)，(234)生成的4个3元循环群.

第六节　正规子群　商群

一、正规子群

一个群G的子群H的左陪集aH未必等于右陪集Ha，但是群的正规子群有这样的性质：任意$a \in G$，$aH = Ha$. 正规子群是非常重要的子群，在讨论群时，几乎处处需要它.

定义1　设H是群G的一个子群，如果对于G的每一个元a，都有

$$aH = Ha,$$

H就叫做G的一个正规子群(不变子群)，记作$H \lhd G$. 正规子群H的左(右)陪集就叫做H的陪集.

群G及单位元群显然都是G的正规子群，只有自身及单位元群为正规子群的群，叫做单群. 交换群G的子群是G的正规子群.

例1　S_3的子群$H_1 = \{(1), (123), (132)\}$，$H_2 = \{(1), (13)\}$，这时

$$(123)H_1 = (1)H_1 = (132)H_1 = H_1(123) = H_1(1) = H_1(132),$$

$$(12)H_1 = (13)H_1 = (23)H_1 = H_1(12) = H_1(23) = H_1(13),$$

故H_1是S_3的正规子群.

又　$(12)H_2 = \{(12), (132)\}$，　$H_2(12) = \{(12), (123)\}$，

因为$(12)H_2 \neq H_2(12)$，故H_2不是S_3的正规子群.

定理1　设H是群G的子群，则以下命题是互相等价的：

(1) 任意$a \in G$，有$aH = Ha$；

(2) 任意$a \in G$，任意$h \in H$，有$aha^{-1} \in H$；

(3) 任意$a \in G$，有$aHa^{-1} \subseteq H$；

(4) 任意$a \in G$，有$aHa^{-1} = H$.

证　(1)⇒(2)：任意$a \in G$，任意$h \in H$，由$aH = Ha$，有$ah \in Ha$，则存在$h_1 \in H$，使得$ah = h_1 a$. 即$aha^{-1} = h_1 \in H$，于是(2)成立.

(2)⇒(3)：$aha^{-1} \in H$，因为 $aha^{-1} \in aHa^{-1}$，故 $aHa^{-1} \subseteq H$，(3)成立.

(3)⇒(4)：由 $aHa^{-1} \subseteq H$ 对任何 G 的元 a 成立，那么对于 a^{-1} 也有 $a^{-1}H(a^{-1})^{-1} \subseteq H$，即 $a^{-1}Ha \subseteq H$，故任意 $h \in H$，有 $a^{-1}ha = h_1$，$h_1 \in H$，所以 $h = ah_1a^{-1} \in aHa^{-1}$，得到 $H \subseteq aHa^{-1}$，故 $aHa^{-1} = H$，(4)成立.

(4)⇒(1)：由于 $aHa^{-1} = H$，于是 $(aHa^{-1})a = Ha$，即 $Ha = aH$.

由定理 1，要检验一个子群是否是正规子群时，可用 4 个条件中的任何一个. 而通常用条件(2)比较方便，因为它指出元素的性质，比证明两个集合相等相对简单一些，例如下面例 2.

例 2 设 A，B 分别是线性群 $GL(n, \mathbf{R})$ 及特殊线性群 $SL(n, \mathbf{R})$ 中的任意元，由于 $|ABA^{-1}| = 1$，所以 $ABA^{-1} \in SL(n, \mathbf{R})$，故

$$SL(n, \mathbf{R}) \lhd GL(n, \mathbf{R}).$$

正规子群还有以下性质：

定理 2 设 $A \lhd G$，$B \lhd G$，则 $A \cap B \lhd G$，$AB \lhd G$.

证 任意 $g \in G$，$c \in A \cap B$，由定理 1 有 $gcg^{-1} \in A$，$gcg^{-1} \in B$，所以 $gcg^{-1} \in A \cap B$，故 $A \cap B \lhd G$.

先证 AB 是 G 的子群，因为 $A \lhd G$，所以 $AB = BA$，由本章第五节定理 1 知 AB 是 G 的子群.

再证 $AB \lhd G$，任意 $g \in G$，$ab \in AB$，有 $gabg^{-1} = (gag^{-1})(gbg^{-1}) \in AB$，所以 $AB \lhd G$. 定理得证.

二、中心与正规化子

设 G 是一个群，把与 G 中所有元素都可交换的元素构成的集合用 $Z(G)$ 表示，即

$$Z(G) = \{a \mid a \in G, \text{任意 } x \in G \text{ 有 } ax = xa\}.$$

显然 G 的单位元 $e \in Z(G)$，故 $Z(G)$ 非空. 又因任意 a，$b \in Z(G)$，任意 $g \in G$，$ag = ga$，$bg = gb$，于是

$$(ab)g = a(bg) = a(gb) = (ag)b = g(ab),$$

所以 $ab \in Z(G)$，再由 $ag = ga$，有 $g = a^{-1}ga$，于是 $ga^{-1} = a^{-1}g$，所以 $a^{-1} \in Z(G)$，故 $Z(G)$ 是 G 的一个子群，同时 $Z(G)$ 是 G 的一个正规子群，$Z(G)$ 叫做群 G 的中心.

当 G 为交换群时，$Z(G) = G$. 当 G 的中心是单位元群时，称 G 没有中心.

由群 G 中心的定义知，若 H 是群 G 的正规子群，那么 G 中任意元均可与 H 交换. 但若 H 不是正规子群，那么 G 中就有与 H 不能交换的元. 更一般地，当 H

是群 G 的一个非空子集时，G 中与 H 能交换的所有元构成的集合，记作$N(H)$，即

$$N(H) = \{a \mid a \in G,\ aH = Ha\}.$$

因为 e 是 G 的单位元，$eH = He$，故 $e \in N(H)$，所以 $N(H)$非空. 若 $a, b \in N(H)$，则 $aH = Ha$，$bH = Hb$，则 $(ab)H = a(bH) = a(Hb) = H(ab)$，故 $ab \in N(H)$. 又由 $aH = Ha$ 有 $aHa^{-1} = H$，即 $a^{-1}H = Ha^{-1}$，所以 $a^{-1} \in N(H)$，故 $N(H)$为 G 的子群. $N(H)$就叫做群 G 中 H 的正规化子.

关于群 G 的正规化子 $N(H)$有以下重要结论.

定理 3　设 H 是群 G 的一个子群，则 H 是其正规化子的正规子群.

证　子群 H 的正规化子为 $N(H) = \{a \mid a \in G,\ aH = Ha\}$，且 $N(H)$是 G 的子群. 而 $H \subseteq N(H)$，H 是 G 的子群，所以 H 是 $N(H)$的子群，再由定理 1 知 H 是 $N(H)$的正规子群. 定理得证.

三、商群

正规子群之所以重要，是因为这种子群的陪集，对于某种与原来的群有密切关系的代数运算来说，也作成一个群，下面来分析.

假如 H 是 G 的正规子群，则 G 关于 H 的陪集的集合用 G/H 表示，即

$$G/H = \{aH \mid a \in G\} = \{Ha \mid a \in G\}.$$

下面来验证 G/H 成为群，它的结合法为子集的乘法.

任意 $aH, bH \in G/H$，由于子集的乘法满足结合律以及 H 是一个正规子群，可得

$$aH \cdot bH = (\{a\}H)(\{b\}H) = \{a\}(H\{b\})H$$
$$= (a(Hb))H = (abH),$$

而 $abH \in G/H$，所以 $aH \cdot bH \in G/H$，所以子集的乘法在 G/H 中是闭合的.

如果把 a 所在的左陪集 aH 用 $\bar{a}$ 表示，因为 G 中元 a，b，c 满足结合律，所以左陪集 $\bar{a}$，$\bar{b}$，$\bar{c}$ 也满足结合律.

G/H 中有单位元 H：任意 $aH \in G/H$，$aH \cdot H = H \cdot aH = aH$，任意 $\bar{a}$ 的逆元$\bar{a}^{-1}$为 $\overline{a^{-1}} = a^{-1}H$，因为 $aH \cdot a^{-1}H = H$.

综上所述，G/H 对于子集的乘法是一个群.

定义 2　设 $H \lhd G$，则 G/H 关于子集乘法构成的群叫做G 关于 H 的商群，当 G 为有限群时，它的阶是$|G : H|$.

例 3　取 S_3 的子群 $H = \{(1), (123), (132)\}$，由于 $|G : H| = 2$，所以 $H \lhd S_3$，而 $S_3/H = \{H, N(12)\}$，其中 $N(12) = \{(12), (13), (23)\}$.

当 G 为加群时，G 中 H 的陪集又叫做 H 的同余类，H 的商群又叫做同余

群，因为它是加群，所以又常常叫做同余加群，有时又叫做差群，常用 $G-H$ 表示. 假如 G 是加群，H 是它的子群，如果 G 中元 a，b 同在 H 的一个同余类中，那么它们的差 $a-b\in H$，用同余式

$$a\equiv b(\bmod H)$$

表示，这时 a，b 又叫做关于 H 同余的. 当 $a\equiv 0(\bmod H)$ 时，a 属于 0 所在的同余类.

例 4 取整数加群 $\mathbf{Z}$ 的子群 $N=\{kn\mid n$ 是固定整数，$k\in\mathbf{Z}\}$，那么 $\mathbf{Z}$ 关于 N 的同余加群 $\mathbf{Z}-N=\{\bar{0},\bar{1},\cdots,\overline{n-1}\}$.

显然交换群的商群仍是交换群，但是它的逆并不成立，即是说有些非交换群的商群也是交换群.

四、换位子群

下面介绍换位子这一概念.

设 a，b 是群 G 中的任意元，如果 $ab=ba$，那么 $a^{-1}b^{-1}ab=e$，如果 $ab\neq ba$，记 $a^{-1}b^{-1}ab=[a,b]$，即有 $ab=ba\cdot[a,b]$. 这就是说 ba 用 $[a,b]$ 右乘后就变成了 ab，这相当于 $[a,b]$ 的作用在于让 a，b 的位子发生了改变. 因此定义 $[a,b]$ 为 a，b 的换位子. 显然，a，b 的换位子是 b，a 的换位子的逆，因为 b，a 的换位子 $[b,a]=b^{-1}a^{-1}ba$.

当 G 为交换群时，由于 $ab=ba$，故只有单位元是它的换位子. 反过来，如果一个群的换位子只有单位元，那么这群一定是交换群. 由于一个群的两个换位子的乘积不再是这样的换位子，因而，一个群的换位子的全体不一定成为群.

定义 3 **G 为一个群，G 中所有换位子生成的群叫做 G 的换位子群，用 $D(G)$ 或 G' 表示.**

定理 4 **群 G 的换位子群 G' 是 G 的正规子群，并且商群 G/G' 是交换群.**

证 G' 是 G 的换位子群，任取 $a\in G'$，$g\in G$，那么 $aga^{-1}g^{-1}\in G'$，再由 G' 是 G 的子群，有 $aga^{-1}=(aga^{-1}g^{-1})g\in G'$，故 G' 为 G 的正规子群. 又任取 aG'，$bG'\in G/G'$，则

$$aG'bG'=(ab)G',\quad bG'aG'=(ba)G'.$$

再由 G' 是 G 的换位子群，知 $(ba)^{-1}ab=a^{-1}b^{-1}ab\in G'$，由本章第五节定理 2 有

$$abG'=baG'.$$

故 $aG'\cdot bG'=bG'\cdot aG'$，于是 G/G' 为交换群.

定理 5 **H 是群 G 的正规子群，那么 G/H 是交换群的充分必要条件是 H 包含 G 的换位子群 G'.**

证 先证必要性：如果 G/H 是交换群，任取 aH, $bH \in G/H$，那么 $(aH)(bH)=(bH)(aH)$，即 $abH=baH$，则有 $h_1 \cdot h_2 \in H$，使得 $abh_1=bah_2$，即 $ab=bah_2h_1^{-1}=bah$，$h \in H$，因为 $a^{-1}b^{-1}ab=h \in H$，而 $a^{-1}b^{-1}ab$ 是属于换位子群 G'的，故 H 包含 G'.

再证充分性：若 H 是包含 G' 的正规子群，因为任意 $a, b \in G$，那么 $a^{-1}b^{-1}ab \in G'$，即 $(ba)^{-1}ab \in G'$，则 $(ba)^{-1}ab \in H$，那么就有 $abH = baH$. 即 $(aH)(bH) = (bH)(aH)$，所以G/H是交换群. 定理得证.

定理 5 说明群的换位子群是使商群成为交换群的最小正规子群.

例 5 求 A_4 的换位子群.

解 $B_4 = \{(1), (12)(34), (13)(24), (14)(23)\}$ 是 A_4 的正规子群，A_4/B_4是交换群，$A'_4 \subseteq B_4$. 由于

$$(123)[(12)(34)](123)^{-1} = (23)(14),$$

因此$((12)(34))$不是 A_4 的正规子群. 同理 B_4 的其余两个 2 阶子群也不是 A_4 的正规子群. 又 A_4 是非交换群，因此 $A'_4 \neq \{e\}$，从而 $A'_4 = B_4$.

第七节 同 构

一、同构

G 与 G_1 是两个群，在很多情况下往往会提出这样的问题：G 与 G_1 是否有相同的构造？什么叫构造相同呢？要回答这个问题必须建立同构的概念.

定义 1 **设 M, M'是两个乘集，σ 是一个 M 到 M'的双射，并且对于 M 中任意两元 a, b，都有**

$$\sigma(a \cdot b) = \sigma(a) \cdot \sigma(b),$$

即任意两元的乘积的像是这两元的像的乘积. 也就是说，当 $a \to \sigma(a)$, $b \to \sigma(b)$ 时，$ab \to \sigma(ab)$，那么 σ 就叫做M 到 M' 上的同构，也称M 与 M' 同构，用 $M \simeq M'$ 表示. 当 $M = M'$ 时，σ 就叫做M 的自同构映射，简称自同构. 恒等映射就是一个自同构.

例 1 设 $H = \{2n \mid n \in \mathbf{Z}\}$，则 H 是群 $\mathbf{Z}$ 的一个子群.

若令 σ: $\mathbf{Z} \to H$ 为 $\sigma(n) = 2n$，则 σ 是 $\mathbf{Z}$ 到 H 的双射，并且对任意 $m, k \in \mathbf{Z}$，

$$\sigma(m+k) = 2(m+k) = 2m + 2k = \sigma(m) + \sigma(k).$$

所以 σ 是 $\mathbf{Z}$ 到 H 的同构映射.

例 2 $S_3=\{(1),(12),(13),(23),(123),(132)\}$，定义 S_3 到自身的一个映射 σ：$(1)\to(1)$，$(12)\to(23)$，$(13)\to(12)$，$(23)\to(13)$，$(123)\to(123)$，$(132)\to(132)$，那么 σ 就是 S_3 的一个自同构映射.

同构关系是一个等价关系，即具有反身性、对称性、传递性.

因为群的同构映射是一个一一映射，故同构的群有相同的阶，群的同构映射还有以下一些重要性质.

定理 1 **设 σ 是群 G 到 G_1 的一个同构映射，则有**

(1) σ 将 G 的单位元 e 映到 G_1 的单位元 e_1；

(2) σ 将 G 中元 a 的逆元 a^{-1} 映到 a 的像 $\sigma(a)$ 的逆元 $(\sigma(a))^{-1}$；

(3) G 中元素 a 与它在 G_1 中的像 $\sigma(a)$ 有相同的阶数.

证 因为 $e^2=e$，于是 $\sigma(e^2)=\sigma(e)$，所以 $\sigma(e)\sigma(e)=\sigma(e)=\sigma(e)e_1$，由消去律，得 $\sigma(e)=e_1$，(1)得证.

因为 $aa^{-1}=e$，所以 $\sigma(aa^{-1})=\sigma(e)$，即 $\sigma(a)\sigma(a^{-1})=e_1$，因而 $\sigma(a^{-1})=(\sigma(a))^{-1}$，(2)得证.

若 $a^k=e$，有 $\sigma(a^k)=(\sigma(a))^k=e_1$，反之也是成立的，所以 a 与 $\sigma(a)$ 有相同的阶数，(3)得证.

根据定义 1，容易验证，如果一个循环群 (a) 是无限群，那么 (a) 就与整数加群 $\mathbf{Z}$ 同构. 如果 (a) 是元数为 n 的有限群，那么它与 $\mathbf{Z}$ 关于 (n) 的同余加群 $\overline{\mathbf{Z}}_n=\mathbf{Z}-(n)$ 同构. 因此任意两个无限循环群都同构，任意两个有限循环群只要元数一致也是同构的.

同构的定义以及性质说明，在同构映射下，对应的元素在各自的运算之下有相同的关系. 因此，要研究一个群，只要找到一个已经研究清楚了的群与其同构，那么这个群也就研究清楚了. 从抽象的观点来看，两个同构的群可以看作是同样的群，同构的重要性也在于此.

定理 2 **(Cayley(卡莱)定理)任意群与它的变换群的子群同构.**

证 考虑群 G 上的变换群 S，建立 G 与 S 的一个子群的同构映射，从而完成定理的证明.

对 $a\in G$，设映射 σ_a：$G\to G$，任意 $x\in G$，$\sigma_a(x)=ax$. 先证明 σ_a 是一个双射. 若 $x,y\in G$，且 $\sigma_a(x)=\sigma_a(y)$，那么 $ax=ay$，即 $x=y$，所以 σ_a 是一个单射. 又任意 $x\in G$，因为 $x=a(a^{-1}x)=\sigma(a^{-1}x)$，故 σ_a 又是满射. 所以 σ_a 是一个双射.

令 $G'=\{\sigma_a\mid a\in G\}$，则 G' 为 S 的一个非空子集，下证 G' 为 S 的一个子群. S 的单位元是 σ_e，这是因为 $\sigma_e(x)=ex=x$，这里的 e 是 G 的单位元. 又若 $\sigma_a,\sigma_b\in G'$，任意 $x\in G$，$\sigma_a\sigma_b(x)=\sigma_a(\sigma_b(x))=\sigma_a(bx)=abx=\sigma_{ab}(x)\in G'$，又设 $\sigma_a\in G'$，$\sigma_{a^{-1}}\cdot\sigma_a=\sigma_e$，$\sigma_a\cdot\sigma_a^{-1}=\sigma_e$，从而 $\sigma_a^{-1}=\sigma_{a^{-1}}\in G'$，故 G' 为 S 的子群.

再证 $G \simeq G'$，考虑映射 $\theta: G \to G'$，对 $a \in G$，$\theta(a) = \sigma_a$，显然 θ 是满射，又若 $\theta(a) = \theta(b)$，从而 $\sigma_a = \sigma_b$，所以 $\sigma_a(e) = \sigma_b(e)$，故 $a = b$，所以 θ 又是单射，故 θ 是双射. 而 $\theta(ab) = \sigma_{ab} = \sigma_a\sigma_b = \theta(a)\theta(b)$，所以 θ 是 G 到 G' 的同构映射. 因此 $G \simeq G'$. 定理得证.

一个集的所有一一变换成为一个变换群，自同构也是一一变换，那么一个乘集的所有自同构是否能够构成群？假如它们成为群，当然这个群就是变换群的子群. 下面对自同构加以讨论.

定理 3 **乘集 M 的所有自同构形成群，叫做 M 的自同构群.**

证 设 σ, τ 是 M 的自同构，对任意 $a \in M$ 有 $\sigma\tau(a) = \sigma(\tau(a))$，所以对于任意 $a_1, a_2 \in M$，有

$$\begin{aligned}\sigma\tau(a_1a_2) &= \sigma(\tau(a_1a_2)) = \sigma(\tau(a_1)\tau(a_2)) \\ &= \sigma(\tau(a_1))\sigma(\tau(a_2)) = \sigma\tau(a_1)\sigma\tau(a_2)\end{aligned}$$

因此 σ, τ 的积 $\sigma\tau$ 也是 M 的自同构.

又因为 $\sigma^{-1}\sigma(a) = \sigma\sigma^{-1}(a) = a$，所以

$$\begin{aligned}\sigma^{-1}(a_1a_2) &= \sigma^{-1}(\sigma\sigma^{-1}(a_1a_2)) \\ &= \sigma^{-1}(\sigma\sigma^{-1}(a_1) \cdot \sigma\sigma^{-1}(a_2)) \\ &= \sigma^{-1}(\sigma\sigma^{-1}(a_1))\sigma^{-1}(\sigma\sigma^{-1}(a_2)) \\ &= \sigma^{-1}(a_1)\sigma^{-1}(a_2),\end{aligned}$$

因此 σ 的逆 σ^{-1} 也是 M 的自同构，因而 M 的所有自同构形成群.

群 G 的自同构群常用 $\mathrm{Aut}(G)$ 表示，恒等自同构是其单位元.

例 3 设 $G = \{e, a\}$ 是一个 2 阶循环群，G 有一个自同构即恒等自同构，且只有一个自同构. 那么 $\mathrm{Aut}(G)$ 是单位元群.

例 4 $G = \{e, a, b, c\}$ 是 4 阶非循环交换群. G 中运算为

$$a^2 = b^2 = c^2 = e, \quad ab = ba = c,$$

$$ac = ca = b, \quad bc = cb = a.$$

计算它的自同构群.

如果 σ 是 G 的一个自同构，那么 σ 必须把 e 映到 e，而把 a, b, c 映到 a, b, c 中的一个，而且各不相同，因此 σ 有以下 6 种可能：

$$\sigma_1: e, a, b, c \to e, a, b, c;$$

$$\sigma_2: e, a, b, c \to e, a, c, b;$$

$$\sigma_3: e, a, b, c \to e, b, a, c;$$

$$\sigma_4: e, a, b, c \to e, b, c, a;$$

$$\sigma_5: e, a, b, c \to e, c, a, b;$$

$$\sigma_6: e, a, b, c \to e, c, b, a.$$

很容易验证这 6 个映射都是 G 的自同构，所以 G 的自同构群一共包含 6 个元素，即

$$\mathrm{Aut}(G) = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\},$$

其中 σ_1 是恒等自同构.

群的自同构是把生成元仍然变成生成元. 因为 $\sigma(a^k) = (\sigma(a))^k$，因而循环群 (a)，若 (a) 是无限群，因为 (a) 只有两个生成元 a 及 a^{-1}，所以它的自同构有两个：

$$\sigma_1: \begin{matrix} a \to a \\ a^{-1} \to a^{-1} \end{matrix}, \qquad \sigma_2: \begin{matrix} a \to a^{-1} \\ a^{-1} \to a \end{matrix},$$

σ_1 即为恒等自同构，因此 (a) 的自同构群是二元循环群. 如果 (a) 是 n 元循环群，因为它有 k 个生成元 a^r，$(r, n) = 1$，$r < n$，所以它有 k 个自同构 $\sigma_r(a) = a^r$. 循环群的自同构群只是交换群，不一定是循环群.

要找出一个已知群的自同构群一般是很困难的，在大多数情况下，群本身的性质不能转移到它的自同构上去.

二、内自同构

下面，再介绍群的一种重要自同构——内自同构.

假定 a 是群 G 中一元，设映射

$$\sigma: g \to g' = aga^{-1}, \ g \in G,$$

验证 σ 是 G 的自同构.

因为任意 $g \in G$，$\sigma(a^{-1}ga) = a(a^{-1}ga)a^{-1} = g$，即 g 的原像是 $a^{-1}ga$，故 σ 是满射. 又若 $aga^{-1} = aha^{-1}$，那么 $a^{-1}(aga^{-1})a = a^{-1}(aha^{-1})a$，即 $g = h$，所以 σ 又是 G 到 G 的单射，故 σ 是双射. 又

$$\sigma: g \to g' = aga^{-1}, \ h \to h' = aha^{-1}, \ g \in G, \ h \in G,$$

则
$$\sigma(gh) = agha^{-1} = (aga^{-1})(aha^{-1}) = \sigma(g)\sigma(h),$$

因此 σ 是 G 的自同构.

定义 2 上述的由一个元 a 决定的自同构 $\sigma: g \to aga^{-1}$，叫做内(自)同构. 其他的自同构，叫做外(自)同构. 元 aga^{-1} 叫做 g 用 a 得到的变形，g 与 aga^{-1} 叫做共轭.

群 G 的所有内同构形成为群，叫做 G 的内同构群.

定理 4　一个群的内同构群是它的自同构群的正规子群.

证　设 σ 是群 G 的任意自同构，τ 是群 G 的内同构，且对 $g \in G$，$\sigma(g) = g_1$，$\tau(g) = aga^{-1}$，于是

$$\sigma\tau\sigma^{-1}(g_1) = \sigma\tau(\sigma^{-1}(g_1)) = \sigma\tau(g) = \sigma(\tau(g)) = \sigma(aga^{-1})$$
$$= \sigma(a)\sigma(g)\sigma(a^{-1}) = \sigma(a)g_1\sigma(a^{-1}) = \sigma(a)g_1(\sigma(a))^{-1},$$

因为 $\sigma(a) \in G$，故 $\sigma(a)g_1(\sigma(a))^{-1} \in G$ 的内同构群，所以定理得证.

一个群没有中心，并且除内同构外没有外同构，叫做完全群，事实上，只要 $n \geqslant 3$ 且 $n \neq 6$，S_n 都是完全群.

定义 3　如果群 G 的子群 H 在群 G 的任一自同构下的像都等于 H，H 就叫做 G 的一个特征子群.

根据特征子群的定义，G 的特征子群在 G 的内同构下的像也等于自己，所以特征子群一定是正规子群. 循环群的子群一定是特征子群. 而正规子群是仅对内同构保持不变的子群，所以正规子群也叫做不变子群. 但是正规子群不一定是特征子群.

三、共轭

共轭也是一个重要的概念，下面加以讨论.

由定义 2 知，a，b 是群 G 中两个元素，如果 G 中有一元素 g，使得 $a = gbg^{-1}$，称 a，b 共轭.

显然，共轭元素有相同的阶. 如果 G 是交换群，由于

$$a = gbg^{-1} = bgg^{-1} = b,$$

故任意元只能够与它自己共轭. 假如 G 中元 a 与它所有的共轭元相等，那么 a 就应在 G 的中心中，这是因为 $a = gag^{-1}$，于是 $ag = ga$. 元素的共轭关系是一个等价关系. 根据此，G 中元素可以分成一些互不相交的共轭元素的集合，叫做共轭类. 在同一共轭类中的元素互相共轭，不同类中的元素彼此不共轭.

例 6　A_4 的 12 个元素分成下列 4 个共轭类：

① $C_1 = \{(1)\}$；

② $C_2 = \{(123), (134), (142), (243)\}$；

③ $C_2 = \{(132), (143), (124), (234)\}$；

④ $C_4 = \{(12)(34), (13)(24), (14)(23)\}$.

$A_4 = C_1 \cup C_2 \cup C_3 \cup C_4$.

设 G 共有 s 个共轭类，记作 C_1，C_2，…，C_s，再设 C_i 中包含 n_i 个元素 ($i = 1, 2, \cdots, s$). 因为

$$G = C_1 \cup C_2 \cup \cdots \cup C_s,\ C_i \cap C_j = \varnothing (i \neq j),$$

所以
$$|G| = n_1 + n_2 + \cdots + n_s.$$

这个式子叫做 G 的群等式或 G 的群方程.

同元素间的共轭关系一样,群 G 的子群也有共轭关系.

定义 4　设 H_1 及 H_2 是群 G 的两个子群,如果存在 G 中一个元素 a 使得

$$H_2 = aHa^{-1},$$

则 H_1 与 H_2 就叫做是共轭的.

例 7　A_4 的 4 个相互共轭的子群有以下五类:

① {(1)};

② {(1), (12)(34)}, {(1), (13)(24)}, {(1), (14)(23)};

③ {(1), (123), (132)}, {(1), (124), (142)}, {(1), (134), (143)}, {(1), (234), (243)};

④ {(1), (12)(34), (13)(24), (14)(23)};

⑤ A_4.

定理 5　H 是群 G 的子群,那么 H 是 G 的正规子群的充分必要条件是 H 与它的所有共轭子群相等.

证　若子群 H 与它的所有共轭子群相等,即 H 只有一个共轭子群,就是它自己.即

$$aHa^{-1} = H \Rightarrow aH = Ha \Rightarrow H \lhd G \Rightarrow aHa^{-1} = H,\text{任意 } a \in H.$$

当然反之也是成立的,因此定理得证.

所以,正规子群也称自共轭子群.

子群的共轭关系是等价关系. G 的子群可以分成一些互不相交的共轭子群的集合,称为共轭子群类.在同一个类中子群是共轭的,同一个共轭子群类中的子群都有相同的阶.例如例 7, A_4 就分成了 5 个共轭子群类.

在同一共轭子群类中共轭的子群的个数有什么规律呢?有以下定理:

定理 6　H 在 G 中的共轭子群的个数等于其正规化子 $N(H)$ 在 G 中的指标 $|G:N(H)|$.

证　$N(H)$ 在 G 中的指标就是 $N(H)$ 的左陪集的个数,因此只要证明:对于 G 中两元素 a 及 b,共轭子群 aHa^{-1} 与 bHb^{-1} 相等的充分必要条件是 a, b 属于 $N(H)$ 的同一左陪集.

容易得到 $aHa^{-1} = bHb^{-1}$ 的充分必要条件是

$$b^{-1}aHa^{-1}b = H,$$

即 $$(a^{-1}b)^{-1}H(a^{-1}b) = H,$$

亦即 $$H(a^{-1}b) = (a^{-1}b)H,$$

故 $$a^{-1}b \in N(H).$$

而这一条件又等价于 a，b 属于 $N(H)$的同一个左陪集. 定理得证.

由定理 6 及 Lagrange 定理知，若 G 的元数是 n，那么 H 的共轭子群的个数是 n 的因数.

由定理 6 可知 H 的全部共轭子群可以由 $N(H)$在 G 中的一个左陪集或右陪集分解给出.

例 8 设 S_4 的子群 $H = \{(1), (12), (34), (12)(34)\}$，可以得到 H 在 S_4 中的共轭子群. 因为

$$N(H) = \{(1), (12), (34), (12)(34), (13)(24), (14)(23), (1324), (1423)\},$$

所以 $|G:N(H)| = 24/8 = 3$，H 共有 3 个共轭子群.

又因为

$$G = N(H) \cup N(H)(13) \cup N(H)(14).$$

所以 H 的 3 个共轭子群为

① H;

② $(13)^{-1}H(13) = \{(1), (14), (23), (14)(23)\}$;

③ $(14)^{-1}H(14) = \{(1), (13), (24), (13)(24)\}$.

四、中心化子

同正规化子的定义一样，可以定义元素的中心化子和子群的中心化子.

定义 5 设 a 是群 G 的一个元素，作 G 的子集

$$Z(a) = \{x \mid xa = ax, x \in G\}.$$

$Z(a)$叫做 a 的中心化子.

很容易验证 $Z(a)$是群 G 的一个子群.

同处理共轭子群的情形一样，G 中元 a 的共轭元的个数等于 a 在 G 内的中心化子 $Z(a)$在 G 中的指标$|G:Z(a)|$.

定义 6 设 H 是群 G 的子集，作 G 的子集 $Z(H) = \{x \mid x \in G, xh = hx, h \in H\}$，即 $Z(H)$是 G 中所有与 H 中任意元可交换的元的集合，$Z(H)$叫做 H 的中心化子.

很显然 $Z(H) = \bigcap_{h \in H} Z(h)$，所以 $Z(H)$也是 G 的子群.

联系子集 H 的正规化子 $N(H)$，有 $Z(H) \subseteq N(H)$，而且 $Z(H)$不一定包

含 H.但当 H 是交换群时 $Z(H)$包含 H.同样 $N(H)$也未必包含 H,但当 H 是子群时,$N(H)$要包含 H.如果 H 是一个子群,那么$Z(H)$是$N(H)$的子群.并且有:

定理 7　设 H 是群G 的一个子群,则$Z(H)$是$N(H)$的一个正规子群.

证　任取 $z \in Z(H)$, $g \in N(H)$,那么对于 H 中任一元素 h,都有

$$(g^{-1}zg)^{-1}h(g^{-1}zg) = g^{-1}z^{-1}ghg^{-1}zg$$
$$= g^{-1}z^{-1}(ghg^{-1})zg.$$

因为 $g \in N(H)$, 所以 $ghg^{-1} \in H$, 因此 z 与 ghg^{-1}可交换,于是

$$(g^{-1}zg)^{-1}h(g^{-1}zg) = g^{-1}ghg^{-1}g = h.$$

由 $Z(H)$的定义知, $g^{-1}zg \in Z(H)$, 因此 $Z(H) \lhd N(H)$.

下面证明一个重要的等式: $xZ(H)x^{-1} = Z(xHx^{-1})$, $x \in G$.

因为 $(xyx^{-1})(xhx^{-1})(xyx^{-1})^{-1} = xyhy^{-1}x^{-1} = xhx^{-1}$, 这里 $y \in Z(H)$, $h \in H$, 所以 $xZ(H)x^{-1} \subseteq Z(xHx^{-1})$.

又设 $z \in Z(xHx^{-1})$,那么 $zxhx^{-1} = xhx^{-1}z$, 因此 $x^{-1}zxh = hx^{-1}zx$, 所以 $x^{-1}zx \in Z(H)$, 即 $z \in xZ(H)x^{-1}$, 那么 $Z(xHz^{-1}) \subseteq xZ(H)x^{-1}$, 于是等式成立.

由该等式得到:

定理 8　正规子群的中心化子是正规子群,共轭子群的中心化子是共轭子群.

因为 H 是正规子群时,等式中 $Z(xHx^{-1}) = Z(H)$, 那就有

$$xZ(H)x^{-1} = Z(H), \text{即 } xZ(H) = Z(H)x,$$

故 $Z(H)$是正规子群,又 xHx^{-1}是 H 的共轭,故共轭子群的中心化子仍是共轭子群.

定理 9　(Sylow **第二定理**)n **元群**G **的** p-Sylow **子群彼此共轭,假如它们的个数是** n_p**,那么** $n_p \mid n$, $n_p = 1(\mathrm{mod}\ p)$.

这是一个非常重要的定理,证明从略.

第八节　同　　态

一、同态

上节学习的同构映射是一个一一映射并且保持运算,所以保证了同构的群有相同的构造.在有些问题上,还会遇到一些保持运算的映射,但不一定是一一

的，下面介绍的同态映射就是这样的.

定义 1 假设 M, M' 是两个乘集，σ 是 M 到 M' 的映射，如果对 M 中任意两个元 a，b，都有

$$\sigma(a \cdot b) = \sigma(a)\sigma(b),$$

那么 σ 就叫做 M 到 M' 的一个同态映射. 如果 σ 是 M 到 M' 内的映射，σ 就叫做是 M 到 M' 内的同态. 如果 σ 是 M 到 M' 上的映射，σ 就叫做是 M 到 M' 上的同态，又称 M 与 M' 同态，用记号 $M \sim M'$ 表示. 如果 $M' \subseteq M$，即 $\sigma(M) \subseteq M$，σ 就叫做 M 的自同态.

显然，当 σ 是双射时，σ 就是同构，因此同构是同态的特例，所以凡是同态具备的性质，同构也是成立的. 由于同态不适合对称律，故同态不是等价关系，因而 $G \sim H$ 与 $H \sim G$ 是不同的.

例 1 设 $G_1 = \{1, -1\}$，则 G_1 对数的普通乘法构成一个群，令 σ: $S_3 \to G_1$ 为

$$\tau \to \begin{cases} 1, & \text{当 } \tau \text{ 为偶置换时；} \\ -1, & \text{当 } \tau \text{ 为奇置换时，} \end{cases}$$

则 σ 是 S_3 到 G_1 的一个同态映射，且 $S_3 \sim G_1$.

群的同态映射有以下一些重要性质.

定理 1 设 σ 是群 G 到 H 的一个同态映射，则有

(1) σ 将 G 的单位元 e 映到 H 上的单位元 e_1；

(2) σ 将 G 中元素 a 的逆元 a^{-1} 映到 $\sigma(a)$ 的逆元，即

$$\sigma(a^{-1}) = (\sigma(a))^{-1};$$

(3) $\sigma(a)$ 的阶数是 a 的阶数的一个因数.

证 (1)，(2)两条性质可以与前一节关于同构映射的相同性质一样地证明，现证明(3).

设 a 的阶数为 k，则 $a^k = e$，于是 $\sigma(a^k) = (\sigma(a))^k = \sigma(e) = e_1$，故 $\sigma(a)$ 的阶数是 k 的一个因数.

如果把一个群的每个元都与单位元对应，那就得到群到单位元群上的同态，这同态叫做零同态.

定理 2 设群 G 与乘集 G' 同态，那么 G' 成群，即一个群的同态像也是群.

证 设 a_1，b_1，$c_1 \in G'$，a，b，$c \in G$，使得

$$\sigma: a \to a_1,\ b \to b_1,\ c \to G_1,\ \sigma \text{ 为同态映射.}$$

因为 G 是群，故 $(a \cdot b) \cdot c = a(b \cdot c)$，于是 $(a_1 \cdot b_1)c_1 = a_1(b_1 \cdot c_1)$，故 G' 中结合律成立. 又设 G 中单位元 e，σ：$e \to e_1$，e_1 为 G' 的元. 那么 $\sigma(ea) = e_1a_1$，又 ea

$= a$, $\sigma(ea) = a_1$, 所以 $e_1 a_1 = a_1$, 故 G' 中有单位元 e_1, $aa^{-1} = e$, 有 $\sigma(aa^{-1}) = e_1$, 则 $\sigma(a) \cdot \sigma(a^{-1}) = e_1$ 即 $a_1 \cdot \sigma(a^{-1}) = e_1$, 故 a_1 的逆元为$\sigma(a^{-1})$.

综上所述,G'是一个群.

由定理 2 的证明可以看到与定理 1 相同的结果,而且群的同态把子群变成子群,正规子群变成正规子群,群的生成元变成生成元.

二、同态核

定义 2 设 σ 是群 G 到 H 的一个同态映射,e'是 H 的单位元,e'的逆像 a 的集合 $E = \{a \in G, \mid \sigma(a) = e'\}$ (E 叫做 e'在 G 的完全像源),那么 E 就叫做同态 σ 的同态核,记为 $\ker\sigma$.

定理 3 假定 G 与 H 是两个群,并且 G 与 H 同态,那么这个同态映射的核 E 是 G 的一个正规子群.

证 设 σ 是 G 到 H 的同态映射,设 a, b 是 E 中任意两元,g 是 G 中的元,那么在 σ 之下 $\sigma(a) = e'$, $\sigma(b) = e'$, e' 为 H 的单位元. 故

$$\sigma(ab^{-1}) = \sigma(a)\sigma(b^{-1}) = e'(\sigma(b))^{-1} = e'e'^{-1} = e',$$

于是 $ab^{-1} \in E$, 因此 E 是 G 的一个子群.

又设 σ: $g \to g_1$, $g_1 \in H$, 则

$$\sigma(gag^{-1}) = \sigma(g)\sigma(a)\sigma(g^{-1}) = g_1 e'(\sigma(g))^{-1} = g_1 e' g_1^{-1} = e',$$

故 $gag^{-1} \in E$, 这即是说 E 是 G 的一个正规子群.

对于任意元的完全像源,有以下定理.

定理 4 假定群 G 与 H 同态,E 是同态核,G 中元 a 在 H 中的像为 a',那么 a'的完全像源是左陪集 aE.

证 左陪集 aE 中的任意元的像都是 $a'e' = a'$, 所以 aE 中的任意元都是 a'的像源,又设 b 的像为 a',那么在同态映射 σ 之下

$$\sigma(a^{-1}b) = (a')^{-1} \cdot a' = e',$$

故 $a^{-1}b \in E$, 所以 $b \in aE$, 即 b 在左陪集 aE 中. 于是 a'的完全像源是 aE.

由定理 3 可知,从一个同态出发就可以得到一个正规子群,也就是它的同态核. 那么从群 G 的一个正规子群出发,能否得到 G 的一个同态像?有以下定理:

定理 5 假定 H 是 G 的正规子群,那么 G 与它关于 H 的商群 G/H 同态.

证 设有 G 到 G/H 的映射 σ: $g \to gH$.

显然 σ 是映上的,为了证明 σ 是一个同态映射,只要证明 σ 保持运算就可以了. 对于 G 中任意两个元素 g_1, g_2,有

$$\sigma(g_1g_2) = g_1g_2H = g_1H \cdot g_2H = \sigma(g_1)\sigma(g_2).$$

因此 σ 保持运算，是一个同态映射.

而且，同态核就是 H，因为核是 G/H 的单位元的原像，又 G/H 的单位元是 H，而且 $\sigma(g) = H \Leftrightarrow gH = H \Leftrightarrow g \in H$，故同态核就是 H.

定理 5 规定的同态映射 σ 叫做 G 到商群 G/H 的自然同态.

以上证明了 G 的每个正规子群都可以作为某个同态映射的核，而且每个商群都是 G 的同态像. 下面来证明 G 的每个同态像都可这样得到：G 的每个同态像都与 G 的某一个商群同构，有下面的同态基本定理.

定理 6　设群 G 与 G' 同态，E 是同态核，那么 $G' \simeq G/E$.

证　因为 $G \sim G'$，假定这时的同态映射是 σ: $a \to a'$，由定理 4 知，a' 的完全像源是 a 所在的左陪集 aE. 我们用 $\bar{a}$ 表示左陪集 aE，命 $\bar{a}$ 与 a' 对应，即 $\bar{a} \to a'$. 因为 $a' = b'$ 时，a, b 同属 E 的一个左陪集，即 $\bar{a} = \bar{b}$，所以这映射就是 G/E 射到 G' 的双射，又因为

$$\bar{a}\,\bar{b} = \overline{ab} \to (ab)' = a'b',$$

即这映射保持运算，所以 $G/E \simeq G'$. 定理得证.

下面给出关于正规子群的一个重要定理.

定理 7　设 A 是群 G 的一个子群，H 是群 G 的一个正规子群，则

$$AH/H \simeq A/A \cap H.$$

证　由于 $H \lhd G$，因此 $H \lhd AH$，令

$$\sigma: A \to AH/H,$$
$$a \to aH$$

在 AH/H 中任取 $(ah)H = aH$，则 $\sigma(a) = aH = (ah)H$，因此 σ 是满射. 对于任意 $a_1, a_2 \in A$，有

$$\sigma(a_1a_2) = a_1a_2H = (a_1H)(a_2H) = \sigma(a_1)\sigma(a_2),$$

因此 σ 是 A 到 AH/H 的一个满同态. 而且对每个 $a \in A$，

$$a \in \ker\sigma \Leftrightarrow \sigma(a) = aH = H(H\text{ 是 }AH/H\text{ 的单位元})$$
$$\Leftrightarrow a \in H \Leftrightarrow a \in A \cap H,$$

即 $\ker\sigma = A \cap H$，由定理 6 知，$AH/H \simeq A/A \cap H$. 定理得证.

习　题　4-1

1. 证明：下面 4 个矩阵对于乘法成群：

$$\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \begin{pmatrix}-1&0\\0&-1\end{pmatrix},\ \begin{pmatrix}1&0\\0&-1\end{pmatrix},\ \begin{pmatrix}-1&0\\0&1\end{pmatrix}.$$

2. 令 $G=\{$一切定义在实数集 $\mathbf{R}$ 上的实函数$\}$,任取 $f, g \in G$,定义 $f+g$ 为:

$$(f+g)(x) = f(x)+g(x),\ \forall x \in \mathbf{R}.$$

证明:G 对于以上代数运算构成一个群.

3. 证明:如果群 G 的每一个元都适合方程 $x^2 = e$,其中 e 是 G 的单位元,那么,G 是交换群.

4. 假定 G 是非空集合,它有一个叫做除法的闭合的结合法a/b,并且

(1) $a/a = b/b = e$, (2) $a/(b/b) = a$, (3) $(a/c)/(b/c) = a/b$.

证明:G 对乘法 $ab = a/b^{-1}$, $b^{-1} = e/b$ 成群,其中 $a/a = e$ 是单位元,$e/a = a^{-1}$ 是 a 的逆元.

5. 设 $f = \begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix}$, $g = \begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix}$,计算下列各式:

(1) $f\cdot g$; (2) $g\cdot f$; (3) f^{-1}; (4) g^{-1}; (5) $(f\cdot g)^{-1}$.

6. 将下列 S_6 中的每一置换写成不相交的循环置换的积.

(1) $\begin{pmatrix}1&2&3&4&5&6\\3&5&6&4&2&1\end{pmatrix}$; (2) $(12)(13)(14)$; (3) $(123)(456)$.

7. 假定元 a 的阶数是 n,证明:a^m 的阶数是$\dfrac{n}{(n,\ m)}$.

8. 设 a, b 是群 G 中的两个元素,求证:

(1) a, a^{-1}, $a^{-1}ba$ 有相同的阶;

(2) ab 与 ba 有相同的阶.

9. 证明:S_n 可以用 $n-1$ 个对换(12), (13), $\cdots$, $(1n)$ $(n>1)$ 生成.

10. 设 $f = (i_1 i_2 \cdots i_k)$ 是一个 k 阶循环置换,证明:f 的阶为 k.

11. 设 G 是一个交换群,m 是一个固定的整数,令 $H = \{a \mid a \in G$, 并且 $a^m = e\}$,证明:H 是 G 的一个子群.

12. 设 A, B 都是群 G 的子群,证明:当且仅当 $AB = BA$ 时,AB 是 G 的子群.

13. 证明:整数加群与由 3 和 4 生成的子群相同.

14. 试求循环加群 $\mathbf{Z}$—(100)的所有子群.

15. 写出 S_3 的子群 $H = \{(1), (13)\}$ 的全体左陪集和全体右陪集.

16. 设 H_1, H_2 是群 G 的两个子群,证明:$H_1 \cup H_2$ 是 G 的子群当且仅当 $H_1 \subseteq H_2$ 或 $H_2 \subseteq H_1$.

17. 假定 H 是 G 的子群,K 是 H 的子群,证明:

$$|G:K| = |G:H|\cdot|H:K|.$$

18. 设 H_1, H_2, N 是群 G 的正规子群,且 $H_2 \subseteq H_2$,则 H_1N 是H_2N的正规子群.

19. 证明:群 G 中子群 H 是正规子群的充分必要条件是:H 的任意两个左陪集的乘积仍然是 H 的一个左陪集.

20. 设 N 是含于群 G 的中心 $Z(G)$ 的子群，则 N 是 G 的正规子群；如果 G/N 是循环群，那么 G 是交换群，试加以证明.

21. 设 A，B 都是群 G 的正规子群，并设它们的阶互素，证明：A 中任一元素与 B 中任一元素都可交换.

22. 证明：群 G 的循环正规子群的子群仍是 G 的正规子群.

23. 证明：包含换位子群的子群是正规子群.

24. 证明：群的中心是特征子群，换位子群是特征子群.

25. 假定群 G 与群 $\overline{G}$ 同态，$\overline{N}$ 是 $\overline{G}$ 的一个正规子群，N 是 $\overline{N}$ 的逆像，证明：$G/N \simeq \overline{G}/\overline{N}$.

26. 证明：非交换单群与它的内同构群同构.

27. 设 G 是一个群，对 $a \in G$，定义左平移 a_l：$G \to G$，对 $x \in G$，$a_l(x) = ax$. 证明：映射 θ：$G \to G_l$，对 $a \in G$，有 $\theta(a) = (a^{-1})_l$（其中 $G_l = \{a_l \mid a \in G\}$）是同构映射.

28. 设 H 和 K 是群 G 的子群，且对任一 $k \in K$，有 $kH = Hk$，证明：$H \triangleleft KH$，$H \cap K \triangleleft K$ 且 $KH/H \simeq K/K \cap H$.

29. 证明：单群的同态像是单群或者单位元群.

30. 假定群 G 只有两个不同的正规子群，G 恰有三个不同的正规子群，G 和 G 同态吗？

31. 设 $G = \{-1, 1\}$，二元结合法是数的乘法，则 G 是一个群. 映射 σ：$S_n \to G$，对 $\alpha \in S_n$，若 α 为偶置换，则 $\sigma(\alpha) = 1$，否则 $\sigma(\alpha) = -1$. 试证 σ 是同态. 那么 $\mathrm{Ker}\sigma$ 是什么？$S_n/\mathrm{Ker}\sigma$ 的元素是什么？

32. 证明：子群 H 是正规子群当且仅当 H 的共轭是 H 自身.

33. G 是一个群. 证明：G 是交换群的充分必要条件是 G 到自身的映射 σ：$a \to a^{-1}$，$a \in G$，是 G 的一个自同构.

第二章 环 和 域

在上一章里，讨论了群的概念及其基本性质，了解到群是只具有一个代数运算的代数系. 在这一章里，要研究具有两个代数运算的代数系，这就是环和域. 在中学代数里已经知道，数可以分成：整数、有理数、实数和复数等，在高等代数里又知道，全体整数是一个数环，而全体有理数，全体实数和全体复数都是一个数域. 由此可见，研究环与域这两个代数系，讨论其基本的性质，并认识几种重要的环和域，一方面对于中学代数有更清楚的了解，另一方面可以得到做进一步研究与学习的基本知识.

第一节 环 的 概 念

一、环的定义

在群里，任意两元只有一个结合法，现在来讨论有两种结合法的代数系，环是其中最基本的.

定义 1 **设 R 是一个非空集合，如果它有两个代数运算，一个叫做加法(用符号"+"表示)，一个叫做乘法(用符号"·"表示)，并且满足下列条件：**

(1) 对于加法成为交换群，叫做 R 的加群；

(2) 对于乘法成为半群，叫做 R 的半群；

(3) 对于加法及乘法适合分配律，即 R 中任意三元 a, b, c，有

$$a(b+c)=ab+ac,\quad (b+c)a=ba+ca,$$

则 R 对于这两个代数运算叫做一个环.

若一个环满足乘法的交换律，即 $ab=ba$，就叫做交换环，否则叫做非交换环. 同群的情况一样，元数是有穷的环，叫做有限环，否则叫做无限环. 下面看一些环的例子.

例 1 全体整数对于数的加法及乘法作成环；全体有理数、全体实数、全体复数对于数的普通加法及普通乘法都作成环.

例 2 令 $\mathbf{Z}[\mathrm{i}]=\{a+b\mathrm{i}\mid a,b\in\mathbf{Z},\ \mathrm{i}^2=-1\}$，$\mathbf{Z}[\mathrm{i}]$对于数的加法及乘法运算是一个环，它是一个数环，通常叫做Gauss 整数环.

例 3 用整数组成的所有 n 阶矩阵(n 是固定的)对于矩阵的加法及乘法成

为一个环，它不是交换环. 如果 R 是环，所有用 R 中元组成的 n 阶矩阵成为一个非交换环，叫做 R 上的n 级全矩阵环，用 R_n 表示.

例 4 设 R 是环，$G=\{u_1, u_2, \cdots, u_n\}$ 是群，结合法是乘法，对于形如

$$\sum_{i=1}^{n} a_i u_i = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n,\ a_i \in R$$

的元规定

$$\sum_{i=1}^{n} a_i u_i = \sum_{i=1}^{n} b_i u_i,\text{当 } a_i = b_i,\ i=1, \cdots, n;$$

$$\sum_{i=1}^{n} a_i u_i + \sum_{i=1}^{n} b_i u_i = \sum_{i=1}^{n} (a_i + b_i) u_i;$$

$$\left(\sum_{i=1}^{n} a_i u_i\right)\left(\sum_{i=1}^{n} b_i u_i\right) = \sum_{i,j=1}^{n} a_i b_j u_i u_j.$$

则所有这样的元形成为环，叫做 G 关于 R 的群环，用$R[G]$表示.

例 5 任意加群，如果规定其中任意两元的乘积为 0(加群的单位元)，那它显然成为环，这样的环叫做零环. 只由一个零元组成的环是零环，有时叫做零.

因为环是一个加群，所以在环中用加法表示的各种性质，也就是加群的各种性质，可以直接得到，主要有以下几条：

① $0+a=a+0=a$;

② $-a+a=a-a=0$;

③ $-(-a)=a$;

④ $a+c=b \rightleftharpoons c=b-a$;

⑤ $-(a+b)=-a-b,\ -(a-b)=-a+b$.

下面主要讨论与乘法有关的各种性质.

(1) $0\cdot a = a\cdot 0 = 0$　　(0 是 R 的零元);

(2) $(-a)b = a\cdot(-b) = -ab,\ (-a)(-b)=ab$;

(3) $a(b-c)=ab-ac,\ (b-c)a=ba-ca$.

证 (1) 由分配律　$0\cdot a=(0+0)a=0\cdot a+0\cdot a$，所以

$$-0\cdot a+0\cdot a+0\cdot a=-0\cdot a+0\cdot a,$$

即得　$0\cdot a=0$.

(2) $ab+(-a)b=[a+(-a)]b=0\cdot b=0$,

$ab+a(-b)=a[b+(-b)]=a\cdot 0=0$,

由 ab 的负元$-ab$ 的惟一性知

$$-ab=(-a)b=a(-b),$$

又 $$(-a)(-b)=-[a(-b)]=-[-(ab)]=ab.$$

(3) $a(b-c)=a[b+(-c)]=ab+a(-c)=ab-ac,$

$(b-c)a=[b+(-c)]a=ba+(-c)a=ba-ca.$

由上面的讨论知，在一个环 R 中进行加、减、乘运算与普通代数中的数的情况是一样的，但对于乘法，两个元素相乘其先后次序一般不能颠倒，除法（乘法的逆运算）在 R 中不一定能施行，因此乘法的消去律不一定成立.

定义 2　假如 a 是环 R 中元，如果 R 中有一元 $b\neq 0$，使得 $ab=0(ba=0)$，那么 a 就叫做 R 的左(右)零因子.

由定义 2，当 $a\neq 0$ 时，b 是 R 的右零因子，一个元如果既是左零因子，又是右零因子，就叫做零因子. 环 R 中元 a 如果不是 R 的左、右零因子，就叫做正则元. 非零环的零元是当然的零因子，环中除当然零因子外，还可能有其他零因子.

例 6　在整数环 $\mathbf{Z}$ 上的 2 阶全矩阵环 $\mathbf{Z}_2$ 中，非零元 $\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}$，$\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$ 都是零因子，因为

$$\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}=0,\ \begin{pmatrix} 0 & 0 \\ b & -a \end{pmatrix}\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}=0,\ \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}\begin{pmatrix} d & 0 \\ -c & 0 \end{pmatrix}=0.$$

定义 3　如果环 R 中除零元外既没有左零因子，又没有右零因子，那么 R 就叫做无零因子环；交换无零因子环叫做整环.

整数环就是一个整环，高斯整数环也是一个整环.

显然一个环是无零因子环的充分必要条件是环中任意元 $a,\ b$，若 $ab=0$，则有 $a=0$ 或 $b=0$.

定理 1　环成为无零因子环的充分必要条件是 R 的乘法适合消去律：

$$a\neq 0,\ ab=ac\Rightarrow b=c,$$

$$a\neq 0,\ ba=ca\Rightarrow b=c,$$

其中 $a,\ b,\ c\in R$.

证　充分性：设环 R 的乘法适合消去律，任意 $a\in R$，$a\neq 0$，若有 $b\in R$，使得 $ab=0=a\cdot 0$，由消去律知 $b=0$，说明 a 不是左零因子，故 R 不含左零因子，同理 R 不含右零因子.

必要性：设环 R 不含左零因子，若 $a,\ b,\ c\in R$，并且 $a\neq 0$，而 $ab=ac$，则 $a(b-c)=0$，由于 a 不是左零因子，因而 $b-c=0$，即 $b=c$，这说明 R 的乘法适合左消去律，同理可证 R 的乘法也适合右消去律. 定理得证.

环中元 a，若 $a^n = 0$，这里的 n 是正整数，那 a 就叫做幂零元，零元是幂零元，非零的幂零元是零因子. 因此在无零因子环中零元是惟一的幂零元，它没有非零的幂零元.

二、环的单位元、逆元

下面，讨论环中乘法的单位元和逆元.

环对乘法可能只成为半群，所以在环的定义中，并不要求对乘法要有单位元及逆元，但在许多环中这种元往往是存在的.

定义 4　环 R 的元素 e 叫做 R 的左(右)单位元，如果对于 R 中任意元素 a 均有 $ea = a(ae = a)$. 如果 e 既是左单位元，又是右单位元，那么 e 就叫做 R 的单位元，这时 R 叫做有单位元环.

整数环的单位元是 1.

一个环未必有单位元，但如果 R 有单位元，则只能有一个，事实上，如果 e_1 及 e_2 都是 R 的单位元，那么 $e_1 = e_1e_2 = e_2$.

定理 2　(1) 若环 R 有左单位元 e_1，右单位元 e_2，则 $e_1 = e_2$，且 e_1 是 R 的单位元.

(2) 若环 R 有仅且有一个左单位元 e，则 e 是 R 的单位元.

证　(1) $e_1 = e_1e_2 = e_2$，故 e_1 是 R 的单位元.

(2) 设 e 是 R 的惟一的左单位元，于是对于 R 中任意元 a，b 均有

$$(ae - a + e)b = aeb - ab + eb = b,$$

故 $ae - a + e$ 是 R 的一个左单位元，那么 $ae - a + e = e$，即 $ae = a$，故 e 也是 R 的右单位元，所以 e 是 R 的单位元.

一个环可能有多个左单位元而没有右单位元，同样也可能有多个右单位元而没有左单位元. 例如所有形如 $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ 的矩阵形成的环，它就没有右单位元，但有无穷多个左单位元 $\begin{pmatrix} 1 & c \\ 0 & 0 \end{pmatrix}$.

单位元不是零因子，但左(右)单位元就不一定，例如左单位元 $\begin{pmatrix} 1 & c \\ 0 & 0 \end{pmatrix}$ 是右零因子，因为 $\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & c \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

环中元 a，如果 $a^2 = a$，a 就叫做幂等元. 因此左(右)单位元是幂等元. 通常所说的幂等元是异于零的元. 如果幂等元 e 是正则元，那么它就是单位元. 这是因为

$$e^2 = e, \quad e^2 a = ea,$$

即
$$e(ea - a) = 0,$$

所以
$$ea = a.$$

定义 5　设环 R 中有单位元 e，对于 R 中元 a，如果存在元素 $b \in R$，使得 $ba = e(ab = e)$，则 b 叫做 a 的左(右)逆元. 如果元素 b 使得 $ab = ba = e$，则 b 叫做 a 的逆元，此时 b 记为 a^{-1}，并且 a 就叫做 R 的一个可逆元.

如果 a 有逆元 a^{-1}，那么 a^{-1} 的逆元就是 a；又若 b 有逆元 b^{-1}，则 ab 的逆元是 $b^{-1}a^{-1}$，即

$$(a^{-1})^{-1} = a, \quad (ab)^{-1} = b^{-1}a^{-1}.$$

零元没有逆元，零因子也没有逆元，这是因为 a 是零因子，那么 $ab = 0$，$b \neq 0$，若 a 有逆元 a^{-1}，则 $a^{-1}ab = a^{-1}0$，即 $b = 0$，与假设不合，故 a 没有逆元. 所以 a 有逆元，就不是零因子.

同单位元的性质相似，有：

定理 3　假定环 R 有单位元 e，$a \in R$，则

(1) 如果 a 有左逆元 b 及右逆元 c，则 $b = c$，且 b 是 a 的逆元；

(2) 如果 a 有且仅有一个左逆元 b，则 b 是 a 的逆元.

证　(1) 由题设知 $ba = ac = e$，于是

$$b = be = b(ac) = (ba)c = ec = c.$$

故 b 是 a 的逆元.

(2) 由于 b 是 a 的惟一的左逆元，则 $ba = e$，那么 $(ab - e + b)a = aba - ea + ba = a(ba) - a + ba = e$，这说明 $ab - e + b$ 是 a 的一个左逆元，则 $ab - e + b = b$，即 $ab = e$，b 又是 a 的右逆元，所以 b 是 a 的逆元.

定理 4　在有单位元的环 R 中，所有可逆元对乘法形成为群 G.

证　设 a，b 是任意的可逆元，则 ab 是可逆元，所以 G 适合群的定义的(1). 由环的定义，G 适合群的定义中的(2)，R 的单位元也就是 G 的乘法的单位元，a^{-1} 就是 a 的逆，故 G 成群.

三、子环

环 R 的子集 A，如果对于 R 的两种代数运算来说作成一个环，A 就叫做 R 的一个子环，R 又叫做 A 的扩张环. 环可以看成是自身的子环，异于自身的子环叫做真子环，任意环都有只由一个零元组成的子环. 同群的情况类似，环中与所有元能够交换的全部元形成子环，叫做环的中心，用 $Z(R)$ 表示，显然交换环的中心就是它自身了.

由定义很容易得到：

环 R 的非空子集 A 作成 R 的一个子环的充分必要条件是对于 A 中任意元素 a，b 均有

$$a-b \in A, \quad ab \in A.$$

第二节　除 环 和 域

一、除环和域

上一节介绍了一个含单位元的环的可逆元的概念，并且看到环里任意一个元不一定是可逆元. 那么在一个环里会不会每一个元都有一个逆元？在一些特殊的情形下这是可能的.

定义 1　一个环 R，如果满足以下条件：

(1) R 至少含有两个元；

(2) R 含有单位元；

(3) R 中每一个非零元都有一个逆元.

R 就叫做一个除环.

定义 2　一个交换的除环叫做一个域.

如果一个域是一个有限集，这个域就叫做有限域.

例 1　有理数集 **Q**，实数集 **R**，复数集 **C** 对于数的普通加法及普通乘法都成为一个域.

例 2　令 p 为一个素数，模 p 的剩余类 $\overline{\mathbf{Z}}_p = \mathbf{Z}-(p)$ 是一个域.

除环和域有以下几个重要性质：

(1) 一个除环没有零因子. 这是因为

$$a \neq 0,\ ab = 0 \Rightarrow a^{-1}ab = b = 0.$$

(2) 一个除环 R 中不为零的元对于乘法来说作成一个群，叫做乘群. 这样一个除环是由两个群：加群和乘群，凑合而成的；而分配律好像是一个纽带，使得这两个群中间发生一种联系.

除环是无零因子环，它的逆一般不成立，但对有限环是成立的.

定理 1　元数大于 1 的有限无零因子环是除环.

证　由第一节的定理 1 知无零因子环中非零元满足乘法的消去律，又由第一章第一节知识，它对除法是闭合的，因此环中所有非零元对乘法成群，所以环成为除环. 定理得证.

由定理 1，还可以得到：元数大于 1 的有限整环是一个域.

因此，有限环如果不是除环，它就应该含有零因子. 在有限环中不是零因子的元都是可逆元，也即任意没有逆元的元都是零因子. 例如有限环没有非零的幂零元，并且只有一个幂等元，那么这环就是除环了.

环成为除环，只要它的所有非零元对乘法能够成为群，也即对于环 R 中任意两元 $a(a \neq 0)$，b，方程 $ax = b$，$ya = b$ 在 R 中有解，那么 R 就是除环. 实际上这两个方程只要一个有解就行了.

定理 2 环 R 成为除环的充分必要条件是：对于 R 中任意两元 a，b，$a \neq 0$，方程 $ax = b$（或者 $ya = b$）在 R 中有解.

证 如果 R 是除环，那么对非零元成为群，当然 $ax = b$（或 $ya = b$）在 R 中有解，故充分性成立. 下证必要性.

设 a，b 为 R 中任意非零的元，$ax = b$ 在 R 中有解，那么 $by = x$ 在 R 中也有解，则有 $aby = ax = b$，于是 $ab \neq 0$，即是说 R 中任意非零元的乘积仍是非零元，因此 R 是无零子环. 又因为 $ae = a$，有 $ae \cdot e = ae$，即 $a(e^2 - e) = 0$，于是 $e^2 = e$，这就是说 e 是无零因子环 R 的幂等元，所以 e 是 R 的单位元.

又 $aa' = e$，知 a' 是 a 的右逆元，但 R 是无零子环，因此 a' 也是 a 的逆元，于是 R 中所有非零的元对乘法成群，因此 R 是除环.

二、子域

定义 2 设 A 为域 F 的一个非空子集，如果它对于 F 的代数运算来说作成一个域，则 A 就叫做 F 的一个子域.

同样可以定义子除环，子整环等.

由定义 2，可以得到域 F 的非空子集 A 作成 F 的一个子域的充分必要条件是：

(1) A 包含不等于零的元；

(2) $a, b \in A \Rightarrow a - b \in A$；

(3) $a, b \in A$，且 $b \neq 0 \Rightarrow ab^{-1} \in A$.

由子域的定义，除环 R 中与所有元能够交换的元的全体形成子域，叫做除环 R 的中心. 除环 R 中所有与子除环 F 中任意元能够交换的元形成除环 R 的子除环，叫做 F 的中心化子，用 $Z(F)$ 表示，即 $Z(F) = \{x \mid x \in R, xy = yx, y \in F\}$.

第三节 同态 同构

一、环的同态 环的同构

由第一章的学习，已经知道了群的同态与同构，为了进一步讨论环的结构，

要把同态、同构推广到环、除环上. 同态、同构是研究环的重要工具.

定义 1 设 R 与 R' 是两个环, σ 是 R 到 R' 的映射, 如果对于 R 中任意元 a, b 均有

$$\sigma(a+b)=\sigma(a)+\sigma(b),\ \sigma(ab)=\sigma(a)\sigma(b),$$

那么映射 σ 就叫做 R 到 R' 的同态, 如果 σ 是 R 到 R' 内的映射, σ 就叫做 R 到 R' 内的同态; 如果 σ 是 R 到 R' 上的映射, σ 就叫做 R 到 R' 上的同态; 这时又说 R 与 R' 同态, 用记号 $R\sim R'$ 表示. 如果 σ 又是单射, 即是可逆的, 它就叫做同构, 这时 R 叫做与 R' 同构, 用 $R\simeq R'$ 表示.

例 1 $\mathbf{Z}$ 是整数, $\mathbf{Z}_n$ 是模 n 的剩余类环, 令

$$\begin{aligned}\sigma:\ & \mathbf{Z}\to\overline{\mathbf{Z}}_n\\ & x\to\bar{x},\ x\in\mathbf{Z}.\end{aligned}$$

易见 σ 是 $\mathbf{Z}$ 到 $\overline{\mathbf{Z}}_n$ 的一个满射. 任取 $x_1, x_2\in\mathbf{Z}$,

$$\sigma(x_1+x_2)=\overline{x_1+x_2}=\overline{x_1}+\overline{x_2}=\sigma(x_1)+\sigma(x_2),$$

$$\sigma(x_1x_2)=\overline{x_1x_2}=\overline{x_1}\ \overline{x_2}=\sigma(x_1)\sigma(x_2),$$

所以 σ 是 $\mathbf{Z}$ 到 $\overline{\mathbf{Z}}_n$ 的同态.

例 2 $K=\{a+b\mathrm{i}\mid, a, b\in\mathbf{Q}\}$ 为高斯数域, $K'=\left\{\begin{pmatrix}a & b\\ -b & a\end{pmatrix}\middle| a, b\in\mathbf{Q}\right\}$ 为除环, 令

$$\begin{aligned}\sigma:\ & K\to K'\\ & a+b\mathrm{i}\to\begin{pmatrix}a & b\\ -b & a\end{pmatrix}.\end{aligned}$$

显然, 这时的 σ 是 K 到 K' 的双射, 又因为

$$(a+b\mathrm{i})+(c+d\mathrm{i})=(a+c)+(b+d)\mathrm{i},$$

$$(a+b\mathrm{i})(c+d\mathrm{i})=(ac-bd)+(ad+bc)\mathrm{i},$$

而
$$\begin{pmatrix}a & b\\ -b & a\end{pmatrix}+\begin{pmatrix}c & d\\ -d & c\end{pmatrix}=\begin{pmatrix}a+c & b+d\\ -(b+d) & a+c\end{pmatrix},$$

$$\begin{pmatrix}a & b\\ -b & a\end{pmatrix}\begin{pmatrix}c & d\\ -d & c\end{pmatrix}=\begin{pmatrix}ac-bd & ad+bc\\ -(ad+bc) & ac-bd\end{pmatrix},$$

故
$$\sigma[(a+b\mathrm{i})+(c+d\mathrm{i})]=\sigma(a+b\mathrm{i})+\sigma(c+d\mathrm{i}),$$

$$\sigma[(a+b\mathrm{i})(c+d\mathrm{i})]=\sigma(a+b\mathrm{i})\cdot\sigma(c+d\mathrm{i}),$$

所以这映射是同构，因此 $K \simeq K'$.

由定义 1 知，环的同构映射是一个等价关系，两个同构的环除了记号外，构造完全一样，它保持了环中用加法、乘法两种运算表示的一切运算性质. 因此若 R 与 R'是两个同构的环，那么 R 是整环，R'也是整环；R 是除环，R'也是除环；R 是域，R'也是域.

环的同态，也保持了环 R 到 R'的加法和乘法两种运算，所以 R 到 R'的同态映射 σ 是 R 的加群到 R'的加群的同态映射，因此 R 的零元在 σ 之下的像为 R'的零元；R 的元 a 的负元 $-a$ 在 σ 之下的像为 $-\sigma(a)$；环的同态把幂零元变成幂零元，幂等元变成幂等元. 同态不是一个等价关系，关于同态，有

定理 1　设 R 和 R'是两个环，且 $R \sim R'$，σ 为同态映射，则有：

(1) 如果 R 是交换环，那么 R'是交换环；

(2) 如果 R 有单位元 e，则 $\sigma(e)$ 为 R'的单位元；

(3) 设 a 为 R 的一个可逆元，其逆元为 a^{-1}，则 $\sigma(a)$ 是 R'的可逆元，并且 $[\sigma(a)]^{-1} = \sigma(a^{-1})$.

证　(1) 设 R 是一个交换环，任取 $a', b' \in R'$，则有 $a, b \in R$，使得 $\sigma(a) = a'$，$\sigma(b) = b'$. 因为 $ab = ba$，所以 $a'b' = \sigma(a)\sigma(b) = \sigma(ab) = \sigma(ba) = \sigma(b)\sigma(a) = b'a'$，即 R'为一个交换环.

(2) 设 e 为 R 的单位元，令 $\sigma(e) = e' \in R'$，任取 $a' \in R'$，由于 σ 是满射，所以存在 $a \in R$，使得 $\sigma(a) = a'$. 因为 $ea = ae = a$，故 $\sigma(e)\sigma(a) = \sigma(a)\sigma(e) = \sigma(a)$，即 $e'a' = a'e' = a'$，故 $\sigma(e) = e'$ 为 R'的单位元.

(3) 设 a 是 R 的一个可逆元，且 $aa^{-1} = a^{-1}a = e$，于是 $\sigma(a)\sigma(a^{-1}) = \sigma(a^{-1})\sigma(a) = \sigma(e)$，故 $\sigma(a)$ 是 R'的一个可逆元，且 $[\sigma(a)]^{-1} = \sigma(a^{-1})$. 定理得证.

要注意的是以上结论反过来却不一定成立. 一个环有没有零因子这一性质经过一个同态映射后是不一定保持的. 也就是说如果 $R \sim R'$，R 是无零因子环，R'中可能有零因子；反过来，假如 R'是无零子环，R 中也可能有零因子，因此 R 是整环时，R'不一定是整环，R'是整环时，R 也不一定是整环. 例如 $\mathbf{Z} \sim \overline{\mathbf{Z}}_6$，$\mathbf{Z}$ 是整环，而 $\overline{\mathbf{Z}}_6$ 有零因子.

同群一样，环也有自同构，自同态整数环和有理数域的自同构都只是恒等同构，即任意元与自己对应的同构，而 $a + bi \to a - bi$ 是复数域的自同构. 由可逆元 a 决定的自同构 $x \to axa^{-1}$，叫作内(自)同构，不是内(自)同构的自同构，叫做外(自)同构. 对于任意内同构不变的子环，叫做不变子环.

群的所有自同构形成自同构群. 对于环也希望能有与这类似的自同态环. 设环 R 的自同态 σ，τ 的积 $\sigma\tau(a) = \sigma(\tau(a))$ 为 R 的自同态；σ，τ 的和 $(\sigma + \tau)(a) = \sigma(a) + \tau(a)$ 却不是 R 的自同态，因为

$$(\sigma+\tau)(ab) \neq (\sigma+\tau)(a)\cdot(\sigma+\tau)(b),$$

故 R 的所有自同态不可能形成为环.

下面将环 R 看成是加群来讨论.

G 是一个加群，σ，τ 为 G 的自同态，因为 $\sigma\tau(a)=\sigma(\tau(a))$，所以 σ，τ 乘积 $\sigma\tau$ 是 G 的自同态，又规定

$$(\sigma+\tau)(a)=\sigma(a)+\tau(a).$$

显然 σ，τ 的和 $\sigma+\tau$ 也是加群 G 的自同态，易证加法的交换律、结合律、分配律均成立，再

$$\sigma(0)=0,\quad -\sigma(a)=\sigma(-a),$$

因此 G 的所有自同态成为有单位元的环，叫做加群 G 的自同态环.

因此，单位元群的自同态环是由零元组成的零环，一般有：

定理 2　元数大于 1 的单纯加群的自同态环是除环.

这是因为由群的同态这一节知识有，单纯加群的异于零的自同态是自同构.

如果把环看成加群，那么它就有自同态环，于是任意环都有自同态环. 需要注意的是，这里所说的自同态是把环看成加群时的自同态，并不是环的自同态.

定理 3　设 R 是有单位元的环（或是无零因子环），a 是其中一元，$\sigma_a(r)=ar$，那么 R 与所有自同态 σ_a 形成的环同构.

证　首先因为映射 $\sigma_a(r)=ar$ 是把 R 看成加群时的自同态，又因为 $a,b\in R$，$\sigma_a(r)=ar$，$\sigma_b(r)=br$，则

$$\sigma_{a+b}(r)=(a+b)r=ar+br=\sigma_a(r)+\sigma_b(r)=(\sigma_a+\sigma_b)(r),$$

$$\sigma_{ab}(r)=(ab)r=(\sigma_a\cdot\sigma_b)r,$$

所以
$$\sigma_{a+b}=\sigma_a+\sigma_b,\quad \sigma_{ab}=\sigma_a\cdot\sigma_b,$$

故这样的自同态 σ_a 形成环 R'. 再定义映射 σ: $R\to R'$ 为 $a\to\sigma_a$，易证 σ 是 R 到 R' 的同态.

当 R 是有单位元或无零因子环时，由于，当 $\sigma_a=\sigma_b$ 时，就有 $ar=br$. 又因为 R 是无零因子环，故有 $a=b$，当 R 有单位元 e 时，由 $ae=be$，也有 $a=b$，于是 σ 又是 R 到 R' 的同构. 定理得证.

由定理 3，就可以找到有单位元的环 R 的同构环 R'.

二、挖补定理

在讨论环的结构时，常常需要从一个已知环出发，构造一个新的环使得它以已知环为子环，碰到这种情况，常要用到下面的关于同构环的定理 4，先证明：

引理 假定集合 S 与 S' 之间存在一个一一映射 σ，并且 S 有加法和乘法，那么可以替 S' 规定乘法和加法，使得 S 与 S' 对于一对加法以及一对乘法来说都是同构的.

证 假定在给定的一一映射之下，S 的元 x 同 S' 的元 x' 对应，规定：

$$a' + b' = c', \quad 若 \quad a + b = c;$$

$$a'b' = d', \quad 若 \quad ab = d.$$

这样规定的法则是 S' 的加法与乘法，因为给了 a' 和 b'，就可以找到惟一的 a 和 b，因而也可找到惟一的 c 和 d，惟一的 c'，d'. 这样规定之后，σ 显然对于一对加法及一对乘法来说都是同构映射.

定理 4 (挖补定理)假定 R 是环 S 的一个子环，R 在 S 里的补集与另一个环 R' 没有共同元，并且 $R \simeq R'$，那么存在一个与 S 同构的环 S'，而且 R' 是 S' 的子环.

证 假设 $R = \{a_r, b_r, \cdots,\}$，$R' = \{a'_r, b'_r, \cdots,\}$，并且在 R 与 R' 之间的同构映射 σ 之下，

$$\sigma: \quad a_r \to a'_r.$$

S 中不属于 R 的元用 $a, b, \cdots$ 表示，这样

$$S = \{a_r, b_r, \cdots, \mid a, b, \cdots,\}.$$

现在将所有的 $a'_r, b'_r, \cdots$ 和所有的 $a, b, \cdots$ 放在一起，作成一个集合 S'，

$$S' = \{a'_r, b'_r, \cdots, \mid a, b\cdots,\},$$

即 $S' = R' \cup (S - R)$，并且规定一个法则

$$\sigma_1: a_r \to a'_r, \ a \to a,$$

即 σ_1 是 S 到 S' 的映射，这对应的关系是 $S - R$ 中元与其自身对应，R 与 R' 中元对应. 那么 σ_1 显然是一个 S 到 S' 的满射. 对于 S 的任意两个不相同的元，这两个元若同时属于 R，或同时属于 $S-R$，那么它们在 σ_1 之下的像显然不同；若是这两个元一个属于 R，一个属于 $S - R$，那么它们在 σ_1 之下的像一个属于 R'，一个属于 $S - R$. 又由于 S' 与 $S - R$ 没有共同元，这两个像也不相同，这样 σ_1 就是 S 到 S' 的一个一一映射.

这样，由引理，可以替 S' 规定加法和乘法使得

$$S \simeq S'.$$

再由 S' 的构成来看，$R' \subset S'$，R' 原来就有加法和乘法，并且是一个环，但还不能说，R' 就是 S' 的子环. 因为由子环的定义，R' 要对于 S' 的运算来说作成一个

子环.

用"$\oplus$"表示 S' 的加法，"$+$"表示 R' 的加法，设 a'_r，b'_r 为 R' 的两个任意元，并且

$$a'_r + b'_r = c'_r.$$

那么由于"$\oplus$"的定义，以及 R 与 R' 同构，有

$$a'_r \oplus b'_r = c'_r, \quad a'_r + b'_r = c'_r.$$

这就是说，假如只看对于 R' 的影响，S' 的加法与 R' 原来的加法没有什么区别，同样可看出，S' 的乘法同 R' 的乘法对于 R' 的影响也是一样的. 因此 R' 就是 S' 的子环. 定理得证.

挖补定理的含义可以用下面的图 4.2.1 来表示.

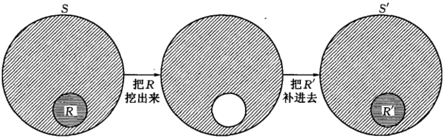


图 4.2.1

第四节　多 项 式 环

一、多项式

这节要介绍一种特殊的环——多项式环，多项式环在环论中占有重要的地位.

假定 R_0 是一个有单位元的交换环，R 是 R_0 的子环，并且包含 R_0 的单位元，$\alpha \in R_0$，那么

$$a_0\alpha^0 + a_1\alpha^1 + \cdots + a_n\alpha^n = a_0 + a_1\alpha^1 + \cdots + a_n\alpha^n \quad (a_i \in R)$$

有意义，是 R_0 的一个元.

定义 1　**一个可以写成**

$$a_0 + a_1\alpha^1 + \cdots + a_n\alpha^n \qquad (a_i \in R,\ n \in \mathbf{N})$$

形式的 R_0 上的元叫做 R 上的关于 α 的一个多项式，$a_i (0 \leqslant i \leqslant n)$ 称为多项式

的系数.

把所有 R 的 α 的多项式放在一起，作成一个集合，这个集合用$R[\alpha]$表示. 并且规定当多项式中 $a_i=0$ 时，$a_i\alpha^i$ 可以略去不写，也就是说，在一个多项式中，可以任意增加或减少系数是零的项. 再规定两个多项式的相等以及它们的加法、乘法，这些也正是普通代数中多项式的计算法则：

(1) $a_0+a_1\alpha+\cdots+a_n\alpha^n=b_0+b_1\alpha+\cdots+b_n\alpha^n$，当 $a_i=b_i$；

(2) $(a_0+a_1\alpha+a_2\alpha^2+\cdots+a_n\alpha^n)+(b_0+b_1\alpha+\cdots+b_n\alpha^n)$
$=(a_0+b_0)+\cdots+(a_n+b_n)\alpha^n$；

(3) $(a_0+\cdots+a_m\alpha^m)\cdot(b_0+\cdots+b_n\alpha^n)=c_0+\cdots+c_{m+n}\alpha^{m+n}$，

其中
$$c_k=a_0b_k+a_1b_{k-1}+\cdots+a_kb_0=\sum_{i+j=k}a_ib_j;$$

这些式子表明，$R[\alpha]$对于加法和乘法都是闭合的，由于还有

$$-(a_0+a_1\alpha+\cdots+a_n\alpha^n)=-a_0-\cdots-a_n\alpha^n\in R[\alpha],$$

所以 $R[\alpha]$是一个环. $R[\alpha]$显然是 R_0 的包括 R 及 α 的最小子环. $R[\alpha]$叫做 R 上的 α 的多项式.

对于一个任意的 α 来说，当系数 $a_1,\ a_2,\ \cdots,\ a_n$ 不全为零时，很可能 α 的多项式

$$a_0+a_1\alpha+\cdots+a_n\alpha^n=0.$$

例如令 $R_0=\mathbf{Q},\ R=\mathbf{Z}$，取 $\alpha=\frac{3}{2}\in\mathbf{Q}$，则 $-2\left(\frac{3}{2}\right)+3=0$.

定义 2 设 R_0 是一个含有单位元的交换环，R 是 R_0 的子环，且 R 含有 R_0 的单位元. x 是 R_0 的一个元. 如果 R 中存在一组不全为零的元$a_0,\ a_1,\ \cdots,\ a_n$ 使得

$$a_0+a_1x+\cdots+a_nx^n=0,$$

x 就叫做 R 上的一个未定元.

在这一节里，主要是讨论未定元的多项式.

对于一个给定的 R_0 来说，R_0 未必含有 R 上的未定元.

例如：R 是整数环，R_0 是包含所有 $a+b\mathrm{i}$ ($a,\ b$ 是整数)的整环，这时对 R_0 的每一个元 $\alpha=a+b\mathrm{i}$ 来说，都有

$$(a^2+b^2)+(-2a)\alpha+\alpha^2=0.$$

二、多项式环

定理 1 如果 R 是有单位元的交换环，则一定存在 R 上的未定元 x，从而 R

上关于未定元 x 的多项式环$R(x)$也存在.

证明 分三步来证明这个定理.

第一步:利用 R 来作一个环 $\bar{P}$,让 $\bar{P}$ 刚好包含所有无穷序列

$$(a_0, a_1, a_2, \cdots),$$

这里 $a_i \in R$,但只有有限个 $a_i \neq 0$,而且规定只在 $a_i = b_i (i = 0, 1, 2, \cdots)$ 时,

$$(a_0, a_1, a_2, \cdots) = (b_0, b_1, b_2, \cdots).$$

规定一个加法:

$$(a_0, a_1, \cdots) + (b_0, b_1, \cdots) = (a_0 + b_0, a_1 + b_1, \cdots).$$

显然这是一个 $\bar{P}$ 的代数运算,而且 $\bar{P}$ 对于这个加法来说作成一个加群.这个加群的零元是$(0, 0, 0, \cdots)$.

再规定一个乘法

$$(a_0, a_1, a_2, \cdots)(b_0, b_1, b_2 \cdots) = (c_0, c_1, c_2, \cdots),$$

这里
$$c_k = \sum_{i+j=k} a_i b_j \qquad (k = 0, 1, 2, \cdots).$$

显然这也是一个 $\bar{P}$ 的代数运算,并且这个乘法适合交换律.

这个乘法也适合结合律,记

$$(a_0, a_1, a_2, \cdots)(b_0, b_1, b_2, \cdots) = (d_0, d_1, d_2, \cdots),$$

$$[(a_0, a_1, \cdots)(b_0, b_1, \cdots)](c_0, c_1, \cdots) = (e_0, e_1, \cdots),$$

那么按照乘法的定义,

$$d_m = \sum_{i+j=m} a_i b_j,$$

$$\begin{aligned} e_n &= \sum_{m+k=n} d_m c_k \\ &= \sum_{m+k=n} \Big(\sum_{i+j=m} a_i b_j \Big) c_k \\ &= \sum_{i+j+k=n} a_i b_j c_k. \end{aligned}$$

把$(a_0, a_1, \cdots)[(b_0, b_1, \cdots)(c_0, c_1, \cdots)]$计算一下,可以得到同样的结果.

这两个代数运算也适合分配律,记

$$(a_0, a_1, \cdots)[(b_0, b_1, \cdots) + (c_0, c_1, \cdots)] = (d_0, d_1, d_2, \cdots),$$

那么,由加法和乘法的定义,

$$d_k = \sum_{i+j=k} a_i(b_j + c_j)$$
$$= \sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j.$$

把 $(a_0, a_1, \cdots)(b_0, b_1, \cdots) + (a_0, a_1, \cdots)(c_0, c_1, \cdots)$ 计算出来,会得到同样的结果.

这样 $\bar{P}$ 作成一个交换环.

在 $\bar{P}$ 里有等式

$$(a_0, 0, 0, \cdots)(b_0, b_1, \cdots) = (a_0 b_0, a_0 b_1, \cdots), \tag{1}$$

由式(1)可以得到

$$(1, 0, 0, \cdots)(b_0, b_1, \cdots) = (b_0, b_1, \cdots),$$

这就是说 $\bar{P}$ 有单位元$(1, 0, 0, \cdots,)$.

第二步:利用 $\bar{P}$ 来得到一个包含 R 的环 P.

由式(1),可以得到

$$(a, 0, 0,\cdots)(b, 0, 0,\cdots) = (ab, 0, 0,\cdots). \tag{2}$$

由加法的定义,我们有

$$(a, 0, 0,\cdots) + (b, 0, 0,\cdots) = (a+b, 0, 0,\cdots). \tag{3}$$

由式(2)和式(3),全体$(a, 0, 0,\cdots)$形成的 $\bar{P}$ 的元作成一个子环 $\bar{R}$,并且

$$(a, 0, 0,\cdots) \to a$$

是 $\bar{R}$ 与 R 间的一个同构映射. 因为 R 同 $\bar{P}$ 没有共同元,由本章第三节定理 4,可以用 R 来代替 $\bar{R}$,而得到一个包含 R 的环 P; P 也是有单位元的交换环,并且 P 的单位元就是 R 的单位元 1.

第三步:最后我们证明 P 包含 R 上的未定元.

记 $$x = (0, 1, 0, 0,\cdots),$$

那么有

$$x^k = (\overbrace{0, \cdots, 0}^{k个}, 1, 0,\cdots). \tag{4}$$

当 $k = 1$ 时,这个式子显然是对的. 假定对于 $k-1$, 式子是对的. 那么,

$$x^k = (0, 1, 0,\cdots)(\overbrace{0, 0, \cdots, 0}^{k-1个}, 1, 0,\cdots)$$

$$= (\sum_{i+j=0} a_i b_j, \sum_{i+j=1} a_i b_j, \cdots),$$

但这里只有 a_1 和 b_{k-1} 等于 1，其余 a_i, b_j 都等于零，所以除了在 $\sum_{i+j=k} a_i b_j$ 这个和里有一项 $a_1 b_{k-1} = 1 \times 1 = 1 \neq 0$ 以外，其余到处都是零，因此就有

$$x^k = (\overbrace{0, 0, \cdots, 0}^{k\text{个}}, 1, 0, \cdots).$$

现在假定在 P 里，

$$a_0 + a_1 x + \cdots + a_n x^n = 0 \qquad (a_i \in R),$$

那么在 $\bar{P}$ 里

$$(a_0, 0, \cdots) + (a_1, 0, \cdots)x + \cdots + (a_n, 0, \cdots)x^n = (0, 0, \cdots).$$

这样，由式(4)和式(1)，

$$(a_0, a_1, \cdots, a_n, 0, \cdots) = (0, 0, \cdots),$$

因而 $$a_0 = a_1 = \cdots = a_n = 0.$$

这正是说，x 是 R 上的未定元. 定理得证.

设 R 是含单位元的交换环，x 是 R 的未定元，R 上关于 x 的多项式环 $R[x]$ 就叫做 R 上一元多项式环：

$$R[x] = \{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in R, 0 \leqslant i \leqslant n, n \text{ 为正整数}\}.$$

$R[x]$ 中的任意两个元素的积与和分别适合式(2)和式(3).

环 $R[x]$ 的元

$$a_0 + a_1 x + \cdots + a_n x^n \tag{5}$$

叫做一元多项式，$a_i (0 \leqslant i \leqslant n)$ 叫做多项式(5)的系数，a_0 为它的常数项. 系数全为零的多项式叫做零多项式，并记为 0. 由未定元的定义可知，$R[x]$ 中每一个非零元可以惟一地表为

$$a_0 + a_1 x + \cdots + a_n x^n \qquad (a_n \neq 0),$$

此时 a_n 叫做它的首系数.

定义 3 **设 $a_0 + a_1 x + \cdots + a_n x^n (a_n \neq 0)$ 是环 R 上关于未定元 x 的多项式，非负整数 n 叫做这个多项式的次数，零多项式没有次数.**

为方便起见，常用符号 $f(x)$ 表示一元多项式环 $R[x]$ 中的元素.

以上所说的概念很容易推广. 一个有单位元的交换环 R_0，和它的一个子环 R，R 包含 R_0 的单位元. 从 R_0 里任意取出 n 个元 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 来，那么可以作

R 上的 α_1 的多项式环$R[\alpha_1]$，然后作$R[\alpha_1]$上的 α_2 的多项式环$R[\alpha_1][\alpha_2]$. 这样下去，可以得到$R[\alpha_1][\alpha_2]\cdots[\alpha_n]$，这个环包括所有可以写成

$$\sum_{i_1 i_2 \cdots, i_n} a_{i_1 i_2 \cdots i_n} \alpha_1^{i_1} \alpha_2^{i_2} \cdots \alpha_n^{i_n}$$

$$(a_{i_1 i_2 \cdots i_n} \in R,\text{但只有有限个 } a_{i_1 i_2 \cdots i_n} \neq 0) \tag{6}$$

形式的元，把具有式(6)的形式的元叫做 R 上的α_1, α_2, $\cdots$, α_n的一个多项式，$a^{i_1 i_2 \cdots i_n}$叫做多项式的系数，环$R[\alpha_1][\alpha_2]\cdots[\alpha_n]$叫做 R 上的α_1, α_2, $\cdots$, α_n的多项式环，这个环也经常用符号$R[\alpha_1, \alpha_2, \cdots, \alpha_n]$表示.

定义 4 R_0 **的** n **个元** x_1, x_2, $\cdots$, x_n **叫做** R **上的无关未定元，如果任何一个** R **上的** x_1, x_2, $\cdots$, x_n **的多项式都不会等于零，除非多项式的所有系数都等于零.**

定理 2 **设** R **是一个有单位元的交换环，**n **是一个正整数，则一定有** R **上的无关未定元**x_1, x_2, $\cdots$, x_n**存在，因此也就有** R **上的多项式环** $R[x_1, x_2, \cdots, x_n]$**存在.**

证 由定理 1，可以找到一个 $R[x_1, x_2, \cdots, x_n]$，使得 x_i 是$R[x_1, x_2, \cdots, x_n]$上的未定元，这样得来的 x_1, x_2, $\cdots$, x_n 是R 上的无关未定元. 这一点用归纳法可以看出：x_1 显然是 R 上的未定元. 假定x_1, x_2, $\cdots$, x_{n-1}是 R 上的无关未定元，由

$$\sum_{i_1 i_2 \cdots i_n} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} = 0 \tag{7}$$

可得
$$\sum_{i_1 i_2 \cdots i_n} (a_{i_1 i_2 \cdots i_n} x_1^{i_1} \cdots x_{n-1}^{i_{n-1}}) x_n^{i_n} = 0,$$

$$\sum_{i_n} \Big(\sum_{i_1 \cdots i_{n-1}} a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_{n-1}^{i_{n-1}} \Big) x_n^{i_n} = 0.$$

这样，因为 x_n 是$R[x_1, x_2, \cdots, x_{n-1}]$上的未定元，式(7)如果成立，一定有

$$\sum_{i_1 \cdots i_{n-1}} a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_{n-1}^{i_{n-1}} = 0 \ (i_n = 0, 1, \cdots).$$

因此，由于 x_1, x_2, $\cdots$, x_{n-1}是 R 上的无关未定元，可得所有

$$a_{i_1 \cdots i_n} = 0.$$

这也就是说，x_1, x_2, $\cdots$, x_n 是 R 上的无关未定元.

上述 n 个无关未定元的多项式的定义与普通 n 个无关变量的多项式的定义并不相同，但这两种多项式有很类似的性质.

三、多项式环的同构与同态

定理 3 设 R 及 $\bar{R}$ 是两个含有单位元的交换环，φ 是 R 到 $\bar{R}$ 的同构映射，$R[x]$与$\bar{R}[x]$分别是 R 及 $\bar{R}$ 上的一元多项式环，则存在 $R[x]$到$\bar{R}[x]$的同构映射 σ，使得$\sigma(x)=x$.

证 设 φ 是 R 到 $\bar{R}$ 的同构映射，即

$$\varphi:\quad R\to\bar{R}$$
$$r\to\bar{r},\qquad \text{任意 } r\in R,$$

设 $f(x)=a_0+a_1x+\cdots+a_nx^n\in R[x]$，作 $R[x]$到$\bar{R}[x]$的映射 σ：

$$a_0+a_1x+\cdots+a_nx^n\to\bar{a}_0+\bar{a}_1x+\cdots+\bar{a}_nx^n.$$

由于 $R[x]$及$\bar{R}[x]$中每个非零多项式的表法是惟一的，以及 φ 是 R 到 $\bar{R}$ 的双射知 σ 是 $R[x]$到$\bar{R}[x]$的双射. 由于多项式环 $R[x]$及$\bar{R}[x]$的加法与乘法分别适合公式(2)，(3)，以及映射 φ 保持 R 到 $\bar{R}$ 的加法及乘法知，如果

$$f(x),\ g(x)\in R[x],\ \sigma(f(x))=\bar{f}(x),\ \sigma(g(x))=\bar{g}(x),$$

这里 $\bar{f}(x)$，$\bar{g}(x)\in\bar{R}[x]$，那么

$$\sigma[f(x)+g(x)]=\overline{f(x)+g(x)}=\bar{f}(x)+\bar{g}(x)=\sigma(f(x))+\sigma(g(x)),$$

$$\sigma(f(x)\cdot g(x))=\overline{f(x)\cdot g(x)}=\bar{f}(x)\cdot\bar{g}(x)=\sigma(f(x))\cdot\sigma(g(x)).$$

故 σ 是 $R[x]$到$\bar{R}[x]$的同构映射，且由 σ 的定义知 $\sigma(x)=x$.

定理 4 设 $R[x_1,x_2,\cdots,x_n]$和 $R[\alpha_1,\alpha_2,\cdots,\alpha_n]$都是有单位元的交换环 R 上的多项式环，$x_1,x_2,\cdots,x_n$是 R 上的无关未定元，$\alpha_1,\alpha_2,\cdots,\alpha_n$是 R 上的任意元，那么 $R[x_1,x_2,\cdots,x_n]$与 $R[\alpha_1,\alpha_2,\cdots,\alpha_n]$同态.

证 用 $f(x_1,x_2,\cdots,x_n)$来表示 $R[x_1,x_2,\cdots,x_n]$的元

$$\sum_{i_1i_2\cdots i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}\cdots x_n^{i_n}.$$

用 $f(\alpha_1,\alpha_2,\cdots,\alpha_n)$来表示 $R[\alpha_1,\alpha_2,\cdots,\alpha_n]$的元

$$\sum_{i_1,\cdots,i_n}a_{i_1\cdots i_n}\alpha_1^{i_1}\cdots\alpha_n^{i_n},$$

那么 $$f(x_1,x_2,\cdots,x_n)\to f(\alpha_1,\alpha_2,\cdots,\alpha_n)$$

是 $R[x_1,x_2,\cdots,x_n]$到 $R[\alpha_1,\alpha_2,\cdots,\alpha_n]$的一个满射，这是因为：给了一个 $R[x_1,\cdots,x_n]$的元 y，由于$x_1x_2\cdots x_n$是无关未定元，只有一种方法可以把 y 写成多项式$f(x_1,x_2,\cdots,x_n)$. 这样，依照规定，y 只有一个像就是 $f(\alpha_1,\alpha_2,\cdots,$

α_n). 另一方面，这个映射显然是一个满射.

由于在 $R[x_1, x_2, \cdots, x_n]$ 或 $R[\alpha_1, \alpha_2, \cdots, \alpha_n]$ 里两个多项式的相加或相乘是适合同一规律的，以上映射是同态映射.

第五节 理 想

一、理想的概念

一个环 R 具有加法、乘法两个代数运算，R 对于其加法正好是一个加群，所以 R 的一个子环 S 是这个加群的正规子群. 由群的知识，可以利用 S 把 R 分成一些陪集，而全体陪集作成商群 R/S. 但必须看到 S 实际上还是 R 的一个子环，因此提出这样的问题：对于商群 R/S，能不能同样借助于环 R 的乘法来规定 R/S 的乘法，进而使 R/S 成为一个环？这样引进理想的概念. 环的理想是一种很重要的子环，这类子环在环论中的地位相当于群论中正规子群的地位.

定义 1　假定 R 是环，S 为 R 的子环，如果对于任意 $r \in R$ 及任意 $a \in S$ 均有 $ar \in S$，$ra \in S$，S 就叫做 R 的一个理想子环，简称为理想.

定理 1　S 是环 R 的非空子集，则 S 是 R 的理想的充分必要条件是

(1) $a, b \in S$，则 $a - b \in S$；

(2) $a \in S$，$r \in R$，则 $ar \in S$，$ra \in S$.

证　必要性是显然的，现证充分性. 由于 $S \neq \varnothing$ 及 S 满足(1)知 S 是 R 的子加群；由 S 满足(2)知 S 是 R 的子环；由(2)及理想的定义知 S 是 R 的理想.

一个环必定含有理想，按定义环 R 及 $\{0\}$ 都是 R 的理想，$\{0\}$ 叫做 R 的零理想，R 本身叫做 R 的单位理想，R 的一个理想，如果不等于零理想和单位理想就叫做 R 的真理想.

有的环除了零理想及单位理想外，没有其他理想，这种环叫做单环，例如除环.

定理 2　一个除环 R 只有两个理想，就是零理想和单位理想.

证　设 R 为除环，A 是 R 的一个理想且 A 不是零理想，那么 $a \in A$，$a \neq 0$，由理想的定义，$a^{-1}a = e \in A$，因而 R 的任意元 $b = b \cdot e \in A$，即 $A = R$.

因此，理想这个概念对于除环或域没有多大用处.

若 N 是环 R 的理想，则 R 关于 N 的同余类 $\bar{a} = a + N$，$\bar{b} = b + N$，可以规定 $\bar{a}$，$\bar{b}$ 的积

$$\bar{a} \cdot \bar{b} = \overline{ab}.$$

显然 R 关于 N 的同余加群 $\overline{R} = R - N$ 对这样规定的乘法是闭合的，再因为 R 中元满足结合律、分配律，所以 N 的同余类也同样满足结合律、分配律，因此 $\overline{R}$ 成

环，这环叫做 R 关于 N 的同余环. $\overline{R}$ 的零元就是 R 的零元 0 所在的同余类 $\overline{0}$，也就是 N 自身，当 R 有单位元 e 时，e 所在的同余类 $\overline{e}$ 就是 $\overline{R}$ 的单位元. 假如 R 是交换环，那么同余环 $\overline{R}$ 也是交换的.

以上介绍的是理想的同余环的概念，下面看一些关于理想的例子.

例 1 $\mathbf{Z}$ 是整数环，$a \in \mathbf{Z}(a \neq 0)$，令

$$S = \{na \mid n \in \mathbf{Z}\},$$

则 S 是 $\mathbf{Z}$ 的理想，特别地，当 $a=2$ 时，则得到全体偶数集 S，S 为 $\mathbf{Z}$ 的一个真理想.

例 2 设 R 是含单位元的交换环，x 是 R 上未定元，则 $R[x]$ 中全体常数项为零的多项式所成的集合 S 是 $R[x]$ 的理想.

证 任取 $f(x) \in S$，则

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x = g(x)x,$$

其中 $g(x) = a_n x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_1$. 设 $f_1(x)$，$f_2(x) \in S$，且

$$f_1(x) = g_1(x)x, \quad f_2(x) = g_2(x)x,$$

这里 $g_1(x)$，$g_2(x) \in R[x]$，那么

$$f_1(x) - f_2(x) = (g_1(x) - g_2(x))x \in S.$$

任取 $h(x) \in R[x]$，$h(x)f(x) = f(x)h(x) = (h(x)g(x))x \in S$，由定理 1 知 S 是 $R[x]$ 的理想.

例 3 设 R 是任意环，a 是 R 的中心 $Z(R)$ 的一个固定元，令

$$A = \{x \in R \mid xa = 0\}，则 A 是 R 的一个理想.$$

证 由 A 的定义知 $0 \in A$，故 $A \neq \varnothing$，又任取 $u, v \in A$，则

$$ua = va = 0,$$

那么
$$(u-v)a = ua - va = 0,$$

所以 $u - v \in A$. 任取 $r \in R$，由于

$$(ru)a = r(ua) = 0,$$

$$(ur)a = u(ra) = u(ar) = (ua)r = 0,$$

所以 $ru, ur \in A$，由定理 1 知 A 是 R 的理想.

同群的情况相似，一个环的任意多个理想的交集仍然是一个理想，但任意两个理想的并集却不一定是理想.

二、主理想

假设 R 是交换环，a 是 R 中的一元，那么 R 中所有形如

$$ra + na\ (r \in R,\ n \text{ 是整数})$$

的元成为一个理想，叫做由元 a 生成的理想，用 (a) 表示. 这是因为

$$(r_1 a + n_1 a) - (r_2 a + n_2 a) = (r_1 - r_2)a + (n_1 - n_2)a \in (a),$$

$$r_1(ra + na) = r_1 ra + nr_1 a = (r_1 r + nr_1)a \in (a),$$

所以 (a) 是理想.

不难看出，任意包含 a 的理想都包含 (a)，故 (a) 也叫做包含 a 的最小理想. 又因为任意多个理想的交集仍是一个理想，所以 (a) 也是所有包含 a 的理想的交集.

如果 R 又有单位元 e，那么 $na = (ne)a$，于是 $ra + na$ 又可以写成

$$(r + ne)a = r'a,\ r' \in R.$$

因此，这时 (a) 是由一切倍元 ra 组成的. 例如整数环 $\mathbf{Z}$ 中理想 (k) 就是由 k 的一切倍数组成的. 要注意的是，由 a 生成的理想 (a) 包含 a，当 R 没有单位元时，所有形如 ra，$r \in R$ 的元虽然也成为理想，但一般它不包含 a，因此它不一定是由 a 生成的理想. 例如 $\overline{\mathbf{Z}}_{12}$ 的子环 $\overline{R} = \{\overline{0}, \overline{2}, \overline{4}, \overline{6}, \overline{8}, \overline{10}\}$ 中，由 $\overline{r}\,\overline{4}$，$\overline{r} \in \overline{R}$ 组成的理想包含 $\overline{4}$，因此它是 $(\overline{4})$，但由 $\overline{r}\,\overline{2}$ 组成的理想不包含 $\overline{2}$，因此它不是 $(\overline{2})$.

由交换环 R 中 n 个元 $a_1, a_2, \cdots, a_n$ 生成的理想 $(a_1, a_2, \cdots, a_n)$ 是所有形如

$$\sum_{i=1}^{n} r_i a_i + \sum_{i=1}^{n} n_i a_i\,(r_i \in R,\ n_i \text{ 为整数})$$

的元组成的，$a_1, a_2, \cdots, a_n$ 叫做它的生成元. 如果 R 有单位元，那么上式可以简写为 $\sum_{i=1}^{n} r_i a_i$ 的形状.

又若 R 是非交换环，a 是 R 中的元，但不在中心 $Z(R)$ 中，那么由 a 生成的理想 (a) 是由所有形如：

$$r_1 a + a r_2 + \sum r_i a r_j + na\ (r_i \in R,\ n \text{ 为整数})$$

的元组成的. 当 R 有单位元时，上式又可简写为 $\sum r_i a r_j$.

环 R 的理想不一定是由一个元生成的，由一个元生成的理想，叫做主理想，零理想是主理想；单位理想 R，当 R 有单位元 e 时是主理想，因为 $R = (e)$.

最后讨论理想与环的关系.

三、理想与环的关系

定理 3 **假定环 R 与环 R' 同态,R 中元 a 在 R' 中的像是 a',那么 R' 的零元 $0'$ 的完全像源 N 是 R 的理想,叫做此同态的同态核. a' 的完全像源是同余类 $a+N$.**

证 首先,因为环的同态是把环看成加群时的同态,所以 N 就是同态核. 由群的同态定理知,N 是加群,并且 a' 的完全像源是 N 的同余类 $a+N$.

其次,当 $a \in N$, $r \in R$ 时,有 $a \to 0'$, $r \to r'$,于是

$$ra \to r'0' = 0', \qquad ar \to 0'r' = 0',$$

所以 ra, $ar \in N$,于是 N 是 R 的理想.

定理 4 **假定 N 是环 R 的理想,那么 R 与同余环 $\overline{R} = R - N$ 同态,同态核就是 N.**

证 设 $a \in R$,让 a 与它所在的 N 的同余类 $\bar{a}$ 对应,即 $a \to \bar{a}$,由群论知这映射是 R 到 $\overline{R}$ 的满射. 又因为

$$a+b \to \overline{a+b} = \bar{a} + \bar{b},$$
$$ab \to \overline{ab} = \bar{a}\,\bar{b},$$

所以 $R \sim \overline{R}$. 由群论知,这个同态核就是 N.

定理 5 **假设环 R 与 R' 同态,同态核是 N, $\overline{R} = R - N$, 那么**

$$R' \simeq \overline{R}.$$

证 因为 $R \sim R'$,假设此时的映射为 $a \to a'$, $\bar{a}$ 是 R 中 a 所在的同余类 $a+N$,由定理 3 知 $\bar{a}$ 是 a' 的完全像源,则可让 $\bar{a}$ 与 a' 对应,即 $\bar{a} \to a'$. 这映射显然是 $\overline{R}$ 到 R' 的双射,又因为

$$\bar{a} + \bar{b} = \overline{a+b} \to (a+b)' = a' + b',$$

$$\bar{a} \cdot \bar{b} = \overline{ab} \to (ab)' = a'b',$$

所以 $\overline{R} \simeq R'$.

因此,任意同态像可以看成同余环,因此理想可以决定环的所有同态. 我们知道,子群、正规子群经过一个同态映射是不变的,子环、理想也是这样.

定理 6 **在环 R 及 R' 的一个同态映射之下,**

(1) R 的一个子环的像是 R' 的一个子环;

(2) R 的一个理想的像是 R' 的一个理想;

(3) R' 的一个子环的逆像是 R 的一个子环;

(4) R'的一个理想的逆像是R的一个理想.

这个定理的证明同群论里的相似定理的证明类似,这里略去.

第六节　极大理想

上节学习了理想的概念及其性质,这一节要介绍极大理想的概念,这是一种比较重要的理想子环.

定义 1　假定 N 是环 R 的理想,如果 $N \subset R$,并且 R 中不存在适合 $N \subset M \subset R$ 的理想 M,那么 N 就叫做 R 的极大理想.

极大理想的定义也可以描述成:N 是环 R 的理想,R 的理想 M 若真包含 N,必有 $M = R$,则 N 叫做 R 的一个极大理想.

由定义 1 知单位理想不是极大理想,单环的零理想是极大理想.

例 1　设 $\mathbf{Z}$ 为整数环,p 是一个素数.则(p)是 $\mathbf{Z}$ 的一个极大理想.

证　$(p) = \{np \mid n \in \mathbf{Z}\}$ 由于 $1 \overline{\in} (p)$,则$(p) \neq \mathbf{Z}$,又设 B 为 Z 的理想,且 B 真包含(p),则存在 $a \in B$,而且 $a \overline{\in} (p)$,由此 $p \nmid a$,故$(p, a) = 1$,则存在$u, v \in \mathbf{Z}$ 使得

$$ua + vp = 1.$$

由此推出 $1 \in B$,则 $B = \mathbf{Z}$,因此,(p)是 $\mathbf{Z}$ 的极大理想.

由例 1 知,一个环可以有很多极大理想.

定理 1　设 $N \neq R$ 是环 R 的理想,同余环 $\overline{R} = R - N$ 除了零理想及单位理想外不再有理想的充分必要条件是 N 为极大理想.

证　设 φ 是 R 到 $\overline{R} = R - N$ 的同态.

充分性:当 N 为极大理想时,设 $\overline{N}$ 为 $\overline{R}$ 的极大理想,并且 $\overline{N} \neq \overline{0}$,那么在 φ 之下 $\overline{N}$ 的逆像 M 是 R 的理想,M 显然包含 N 且不等于 N,所以

$$M = R, \quad \overline{N} = \overline{R}.$$

因此,$\overline{R}$ 只有零理想与单位理想.

必要性:假设 N 不是极大理想,则 $N \subset M \subset R$,M 为 R 的理想,且 M 既不等于 N 也不等于 R,那么在 φ 之下的 M 的像 $\overline{M}$ 是 $\overline{R}$ 的理想,由于 M 大于 N,

$$\overline{M} \neq \overline{0}.$$

$\overline{M}$ 也不会是 $\overline{R}$,否则,对于 R 中任意元 r,可以找到 M 的元 b,使得

$$\bar{r} = \bar{b}, \; r - b \in N \subset M.$$

于是,由于 M 是理想,可以找得到 $r \in M$, $M = R$,与假设不合.这样 $\overline{R}$ 除了零理想及单位理想以外还有理想 $\overline{M}$.定理得证.

一般的，一个域只有零理想及单位理想，反过来，一个只有这两个理想的环不一定是一个域.

定理 2　若有单位元的交换环 R 除了零理想与单位理想以外没有其他的理想，那么 R 一定是一个域.

证　设 a 是 R 中不为 0 的任意元，由 a 生成的主要理想 (a) 显然不是零理想，于是由假定 $(a)=R$，因而 R 的单位元 $e\in(a)=R$，但是 (a) 的元都可以写成 $ra\,(r\in R)$ 的形式，所以

$$e=a'a\quad(a'\in R).$$

这样，R 的每个不等于零的元都有一个逆元，R 是一个域.

由以上两个定理得到

定理 3　假定 R 是一个有单位元的交换环，N 是 R 的极大理想的充分必要条件为同余环 $\overline{R}=R-N$ 是域.

由定理 3，对一个有单位元的交换环 R，只要找到 R 的一个极大理想 N，就可得到一个域 $\overline{R}=R-N$.

由定理 3 及例 1，得到 $\mathbf{Z}-(p)$ 是一个域.

第七节　分　式　域

一、分式域

环之所以不能成为除环或者域，是因为它的元不一定都有逆元. 能否将这些逆元与逆元的乘积都添入其中使它成为除环？例如整数集 $\mathbf{Z}$ 只成为整环，它不能成为除环. 因为它的元除 1 及 -1 外，其余的都没有逆元. 如果把这些逆元以及逆元的乘积也就是所有由整数做成的分式添入，那就成为有理数域 $\mathbf{Q}$. 仿照以上作法，从交换环 R 中做出这些逆元，添加于 R，使它成为包含 R 的域. 也就是说，对于一个交换环，可以做一个域把所给的环嵌入. 分两步来进行，首先假设已知有一个包含 R 的除环 K，在 K 中如何去找这些分式或商，使它们成为包含 R 的域. 其次，假设包含 R 的除环不是已知的，如何从一个交换环中做出分式成为包含它的域.

如果 a，b 为除环中非零的元，且 $ab=ba$，则 $ab^{-1}=b^{-1}a$，即 b 左除 a 与 b 右除 a 结果是一致的，把它叫做 b 除 a 的分式，用 $\frac{a}{b}$ 来表示，即

$$\frac{a}{b}=b^{-1}a=ab^{-1}.$$

定理 1 设 K 是包含交换环 R 的除环，那么 K 中所有的分式

$$\frac{a}{b}\quad (a,\ b\in R,\ b\neq 0)$$

形成一个包含 R 的域 F，这 F 叫做 R 的分式域.

证 首先验证以下计算法则是成立的：

(1) $\frac{a}{b}=\frac{c}{d}$，当 $ad=bc$；

(2) $\frac{ca}{cb}=\frac{ac}{bc}=\frac{a}{b}$；

(3) $\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$；

(4) $\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}$.

这是因为当 $\frac{a}{b}=\frac{c}{d}$ 时，即 $b^{-1}a=cd^{-1}$，那么 $ad=bc$，故(1)成立.

又 $\frac{ca}{cb}=ca(cb)^{-1}=ca\cdot b^{-1}c^{-1}=ab^{-1}=\frac{a}{b}$，同理 $\frac{ac}{bc}=\frac{a}{b}$，故(2)成立.

$\frac{ad+bc}{ba}=(ba)^{-1}(ad+bc)=(ba)^{-1}ad+(ba)^{-1}(bc)=\frac{ad}{ba}+\frac{bc}{ba}=\frac{a}{b}+\frac{c}{d}$，故(3)成立.

$\frac{a}{b}\cdot\frac{c}{d}=b^{-1}a\cdot d^{-1}c=ac\cdot(b^{-1}d^{-1})=ac(bd)^{-1}=\frac{ac}{bd}$，故(4)成立.

而 $0=\frac{0}{a}$，所以 0 也是分式.

因为 $\frac{a}{b}+\frac{-a}{b}=\frac{a+(-a)}{b}=\frac{0}{b}=0$，所以$\frac{a}{b}$的负元是$\frac{-a}{b}$，由以上知，$F$ 是一个加群.

又因为 $a\neq 0$，$\frac{a}{a}=e$（e 为 K 的单位元），所以 e 也是分式.

而 $\frac{a}{b}\cdot\frac{b}{a}=\frac{ab}{ab}=e$，故$\frac{a}{b}$的逆元为$\frac{b}{a}$，于是 F 对乘法成为群. 于是 F 成为除环，显然 F 满足交换律，故 F 为域.

最后，对于 R 中任意元 a，若 b 是 R 中任意的非零元，总是有 $a=\frac{ab}{b}$，故 $a=\frac{ab}{b}\in F$，故$R\subseteq F$. 因此定理得证.

从定理 1 的证明中发现只要 R 能够嵌入到一个除环中，就可得到 R 的分式域. 而且 F 中分式的运算法则完全由 R 的运算惟一确定，所以 F 的构造完全是

由 R 确定的，因此有

定理 2　同构的环的分式域也同构.

另外，从定理 1 的证明中还会得到，如果 K 中还有包含 R 的域 H，那么 H 一定包含 R 的分式域 F，即 $F \subseteq H$，即是说 F 是包含 R 的最小域. 这时因为任取 $x = ab^{-1} = \frac{a}{b} \in F\ (a,\ b \in R,\ b \neq 0)$，那么 $a,\ b \in H$，$b^{-1} \in H$，则 $ab^{-1} \in H$. 所以 $F \subseteq H$.

下面来讨论一下，对于一般的交换环 R（包含它的除环 K 不是已知）是否有分式域？如果有，那么分式域的分式是怎么表示的呢？

另外，还由于域及除环是没有零因子的，如果 R 有分式域 F，那么 F 没有零因子，而 R 是 F 的子集，故 R 也就没有零因子. 这应该是 R 有分式域的一个必要条件，而实际上这个条件是充分的.

二、环与分式域

定理 3　交换环 R 有分式域的充分必要条件是 R 为整环.

证　定理的必要性已经在前面论述了，下面证充分性. 分三步来证明.

第一步：作一个域 $\overline{F}$.

设 $R = \{a,\ b,\ c,\cdots\}$，令集合

$$A = \{\text{全体符号为}\frac{a}{b} \mid a,\ b \in R,\ b \neq 0\},$$

显然 $A \neq \varnothing$.

在 A 中定义关系"$\sim$"：

$$\frac{a}{b} \sim \frac{c}{d} \Leftrightarrow ad = bc,\ bd \neq 0.$$

于是这关系满足反身性、对称性和传递性. 满足传递性是因为由 $\frac{a}{b} \sim \frac{c}{d}$，$\frac{c}{d} \sim \frac{e}{f}$ 知 $ad = bc$，$cf = de$，$bd \neq 0$，$df \neq 0$，有 $adf = bcf = deb$，R 是整环，没有零因子，那么 $af = be$，所以 $\frac{a}{b} \sim \frac{e}{f}$. 因此关系"$\sim$"是一个等价关系. 这个等价关系把 A 分成若干个等价类，令 $\overline{F}$ 为全体等价类的集合，即

$$\overline{F} = \left\{\left[\frac{a}{b}\right] \,\middle|\, \frac{a}{b} \in A\right\}.$$

在 $\overline{F}$ 中规定

$$\left[\frac{a}{b}\right]+\left[\frac{c}{d}\right]=\left[\frac{ad+bc}{bd}\right],$$

$$\left[\frac{a}{b}\right]\cdot\left[\frac{c}{d}\right]=\left[\frac{ac}{bd}\right].$$

上述规定是 $\overline{F}$ 的两个代数运算(叫做加法与乘法),事实上,

(1) 因为 $b\neq 0$, $d\neq 0$,故 $bd\neq 0$,那么 $\left[\frac{ac}{bd}\right]$及$\left[\frac{ad+bc}{bd}\right]$是 $\overline{F}$ 中元素.

(2) 设 $$\left[\frac{a}{b}\right]=\left[\frac{a'}{b'}\right],\left[\frac{c}{d}\right]=\left[\frac{c'}{d'}\right],$$

则 $$ab'=a'b,\ cd'=c'd,\ bb'\neq 0,\ dd'\neq 0,$$

从而 $$ab'dd'=a'bdd',\ cd'bb'=c'dbb',$$

$$(ad+bc)b'd'=(a'd'+b'c')bd,\ b'd'bd\neq 0,$$

故 $$\left[\frac{ad+bc}{bd}\right]=\left[\frac{a'd'+b'c'}{b'd'}\right],$$

另一方面 $$acb'd'=a'c'bd,\ b'd'bd\neq 0,$$

故 $$\left[\frac{ac}{bd}\right]=\left[\frac{a'c'}{b'd'}\right].$$

这就是说两个类的相加与相乘的规定与类的代表选择无关.

由于

(1) $\left[\frac{a}{b}\right]+\left[\frac{c}{d}\right]=\left[\frac{c}{d}\right]+\left[\frac{a}{b}\right]=\left[\frac{ad+bc}{bd}\right]$;

(2) $\left[\frac{a}{b}\right]+\left(\left[\frac{c}{d}\right]+\left[\frac{e}{f}\right]\right)=\left(\left[\frac{a}{b}\right]+\left[\frac{c}{d}\right]\right)+\left[\frac{e}{f}\right]$;

(3) $\left[\frac{a}{b}\right]+\left[\frac{0}{d}\right]=\left[\frac{ad}{bd}\right]=\left[\frac{a}{b}\right]$;

(4) $\left[\frac{a}{b}\right]+\left[\frac{-a}{b}\right]=\left[\frac{0}{b}\right]$,

故 $\overline{F}$ 成为一个加群. $\overline{F}$ 的乘法显然也适合结合律、交换律,并且容易验证加法与乘法适合分配律,所以 $\overline{F}$ 是一个环. 这个环里,$\left[\frac{a}{a}\right]$是单位元,$\left[\frac{a}{b}\right]$的逆元是$\left[\frac{b}{a}\right]$ $(a\neq 0)$,于是 $\overline{F}$ 是一个域.

第二步:作域 F 使 R 为 F 的子环.

令 $\overline{R}=\left\{\left[\frac{ad}{d}\right]\middle| d\neq 0,\ d \text{ 为 } R \text{ 中固定元}, a\in R\right\}$,作 R 到 $\overline{R}$ 的映射 ψ,

$$\psi:\quad R \to \overline{R}$$

$$a \to \left[\frac{ad}{d}\right], \text{任意 } a \in R.$$

易知 ψ 是 R 到 $\overline{R}$ 的双射,并且任取 $a, b \in R$, 有

$$\psi(a+b) = \left[\frac{(a+b)d}{d}\right] = \left[\frac{ad}{d}\right] + \left[\frac{bd}{d}\right] = \psi(a) + \psi(b),$$

$$\psi(ab) = \left[\frac{abd}{d}\right] = \left[\frac{ad}{d}\right]\left[\frac{bd}{d}\right] = \psi(a)\psi(b).$$

于是 $\overline{R}$ 为 $\overline{F}$ 的子环,且 $R \simeq \overline{R}$, ψ 就是 R 到 $\overline{R}$ 的同构映射.

由于 $R \cap \overline{F} = \varphi$, 由挖补定理知,在 $\overline{F}$ 中可以用 R 代替 $\overline{R}$,这样得到了一个以 R 为子环的域 F,并且 $F \simeq \overline{F}$. 假设这个同构映射为 τ,那么 τ 限制在 R 上就是 ψ. 并且规定,在 F 中,元素 $\left[\frac{d}{bd}\right] = \left[\frac{bd}{d}\right]^{-1}$, 记作 b^{-1}(其中 $b \in R$, $b \neq 0$).

第三步:证明 F 是 R 的分式域.

下面证明 F 是所有元 $\frac{a}{b} = ab^{-1} = b^{-1}a$ ($a, b \in R$, $b \neq 0$) 所作成的域. 为此任取 $\left[\frac{a}{b}\right] \in \overline{F}$, 则

$$\left[\frac{a}{b}\right] = \left[\frac{ad}{d}\right]\left[\frac{d}{bd}\right] = \left[\frac{ad}{d}\right]\left[\frac{bd}{d}\right]^{-1},$$

这里 $\left[\frac{ad}{d}\right]$, $\left[\frac{bd}{d}\right] \in \overline{R}$, 这就是说 $\overline{F}$ 中每个元素都可能写成

$$xy^{-1} (x, y \in \overline{R}, y \neq 0).$$

由于 F 与 $\overline{F}$ 之间的同构映射 τ 的作用可知,F 的每一个元素可以写成

$$ab^{-1} = \frac{a}{b} \qquad (a, b \in R, b \neq 0),$$

于是 F 是 R 的分式域.

由定理 3 的证明可以知道把一个整环嵌入到一个域的方法,因此对于任一整环都可以作出它的分式域.

习 题 4-2

1. 假定 R 是实数集,加法"+"是数的普通加法,乘法"×"是

$$a \times b = |a| b,$$

这时 R 是否成环?

2. 在加群 G 中规定乘法如下:

$$g_1 g_2 = 0, \quad 任意\ g_1, g_2 \in G.$$

证明:G 是一个环.

3. 证明:假如 e 是环 R 的左单位元,如果 R 没有左零因子,那么 e 是 R 的单位元.

4. 假定环 R 有单位元,E 是全矩阵环 R_n 的单位元,证明:R_n 的中心

$$Z(R_n) = Z(R)E.$$

5. 证明:在有限环中正则元都是可逆元,也就是说,有限环如果有正则元,它就有单位元.

6. 令 A 是一个含单位元的环.证明:

(1) 若 a 是 A 的一个可逆元,则 $-a$ 也是 A 的一个可逆元,且 $(-a)^{-1} = -a^{-1}$;

(2) 若 a, b 是 A 的两个可逆元,则 ab 也是 A 的可逆元,且 $(ab)^{-1} = b^{-1}a^{-1}$;

(3) A 的可逆元不是右零因子,也不是左零因子.

7. 假如 $n(n > 1)$ 不是素数,那么 $\mathbf{Z} - (n)$ 就不成为域.

8. R 是一个环,且 R 的元数大于 1,对于任意 $0 \neq a \in R$ 均有惟一元素 $b \in R$,使得 $aba = a$,证明:

(1) R 不含零因子;

(2) $bab = b$;

(3) R 有一个单位元;

(4) R 是一个除环.

9. 令 R 是一个除环,$R_P = \{a \in R \mid \forall x \in R,\ xa = ax\}$,证明 R_P 是一个域.

10. 试作由两个元组成的除环.

11. 设 R 是一个环,证明 R 的中心 $Z(R)$ 是 R 的一个可交换的子环.

12. 证明:有理数加群的自同态环与有理数域同构.

13. 证明:所有整数组成的加群与所有偶数组成的加群同构,但整数环不与偶数环同构.

14. 令

$$R = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \middle| a, b, c \in \mathbf{Z} \right\},$$

$$\overline{R} = \left\{ \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} \middle| a, c \in \mathbf{Z} \right\},$$

证明:环 R 与 $\overline{R}$ 同态.

15. 设 R 是一个环,且 R 是交换环,试证:R 的全体幂零元的集合 S 是 R 的一个子环.

16. 用 $\mathbf{Z}_3$ 表示模 3 的剩余类作成的集合,找出加群 $\mathbf{Z}_3$ 的所有自同构映射,再找出域 $\mathbf{Z}_3$ 的所有自同构映射.

17. 任意两个有单位元的 m 元环,如果 m 没有平方数因子,试证这两个环同构.

18. 设 R 是含单位元 $1 \neq 0$ 的可交换环,$R[x]$ 中一个 m 次多项式与一个 n 次多项式的乘

积是不是一个 $m+n$ 次多项式?

19. 设 R 是含单位元 $1\neq 0$ 的整环,$f(x)\in R[x]$,试证:$f(x)$是$R[x]$中可逆元当且仅当 $f(x)\in R$ 且是 R 中可逆元.

20. 假定 R 是模 7 的剩余类环,在 $R[x]$中把乘积

$$([3]x^3+[5]x-[4])([4]x^2-x+[3])$$

计算出来.

21. 证明:

(1) $R[a_1, a_2]=R[a_2, a_1]$;

(2) 若 $x_1, x_2, \cdots, x_n$ 是 R 上的无关未定元,那么每一个 x_i 都是 R 上的未定元.

22. 证明:

(1) 若 $x_1, x_2, \cdots, x_n$ 与 $y_1, y_2, \cdots, y_n$ 是 R 上的两组无关未定元,那么

$$R[x_1, x_2, \cdots, x_n]\simeq R[y_1, y_2, \cdots, y_n];$$

(2) R 上的一元多项式环 $R[x]$能与它的一个真子环同构.

23. 设 F 是一个域,环 $A=\left\{\begin{pmatrix}a & b\\0 & 0\end{pmatrix}\middle| a, b\in F\right\}$,证明:

$$S=\left\{\begin{pmatrix}0 & a\\0 & 0\end{pmatrix}\middle| a\in F\right\}$$

是 A 的理想.

24. $\mathbf{Z}[\mathrm{i}]=\{a+b\mathrm{i}\mid a, b\in\mathbf{Z}, \mathrm{i}^2=-1\}$ 是高斯整数环,证明:

$$A=\{2a+2b\mathrm{i}\mid a, b\in\mathbf{Z}, \mathrm{i}^2=-1\}\text{ 是 }Z[\mathrm{i}]\text{ 的理想.}$$

25. 设 R 是交换环,a 是 R 的元素,证明:$S=\{ra\mid r\in R\}$ 是 R 的理想,问 S 是不是(a)?

26. 证明:假定 R 是无数大于 1 的整环,并且只包含有限个理想,那么 R 是域.

27. 证明:假定 R 是单环,$R^2\neq 0$,e 是幂等元,那么eRe也是单环.

28. 证明:非幂零单环有单位元时中心是体,没有单位元时中心是零.

29. 证明:有单位元的环有极大理想.

30. 假设 $\mathbf{Q}$ 为有理数域,那么 $\mathbf{Q}[x]$中(x)是否为极大理想?

31. 证明:一个域 F 是它自己的分式域.

32. 假设 p 是素数,试证所有形如 $\frac{m}{n}$,$(n, p)=1$ 的有理数集成为整环,并求它的分式域.

33. 证明:假定 R 是一个有两个以上的元的环,F 是一个包含 R 的域,那么 F 包含 R 的一个分式域.

参考文献

1. 张禾瑞. 近世代数基础. 北京:高等教育出版社,1978
2. 熊全淹. 近世代数. 武汉:武汉大学出版社,2004
3. 卢景波,蒋滋梅,刘云英. 抽象代数基础. 北京:北京师范大学出版社,1990
4. 王萼芳. 有限群论基础. 北京:清华大学出版社,2002
5. 丘维声. 抽象代数基础. 北京:高等教育出版社,2003
6. 盛德成. 抽象代数. 北京:科学出版社,2001
7. 刘绍学. 近世代数基础. 北京:高等教育出版社,1999

第五篇　人工神经网络

本篇通过简要介绍人工神经网络的基本理论、模型、算法和应用，使读者掌握人工神经网络的基本知识和应用方法，为将来更深入的学习和研究打下一定的基础. 第一章介绍了人工神经网络的基础知识，第二章对几个基本的人工神经网络模型和算法进行了讨论，第三章分别对人工神经网络在手写体汉字识别、信号处理和自动控制中的应用进行了介绍.

第一章　概　　述

本章对人工神经网络的发展、现状和基础知识进行了简要介绍，包括神经元模型、激活函数以及人工神经网络的表示、结构和学习等基本问题.

第一节　人工神经网络的发展与现状

一、人工神经网络的发展

人工神经网络是模仿人的脑神经结构和功能以及人类思维和处理问题等功能的新型信息处理系统. 由于人工神经网络具有复杂的动力学特性、并行处理机制、学习、联想和记忆等功能以及它的高度自组织、自适应能力和灵活性而受到自然科学领域学者的广泛重视. 人工神经网络的研究始于 1943 年，心理学家 W. S. McCulloch 和数学家 W. Pitts 提出了 MP 模型，该模型除连接权不调整外，其他与现在的阀值单元模型基本相同. MP 模型的提出不仅具有开创意义，而且还为以后的研究工作提供了依据.

1949 年心理学家 D. O. Hell 提出了神经元之间突触强度调整的假设. 他认为学习过程是在突触上发生的，连接权的调整正比于两相连神经元之间激活值的乘积，这就是有名的 Hebb 学习规则. 直到现在，Hebb 学习规则仍然是人工神经网络中的一个极为重要的学习规则.

20 世纪 50 年代末，F. Rosenblatt 提出了著名的感知机(Perception)模型，这是第一个完整的人工神经网络模型. 这个模型由阀值单元构成，初步具备了并行处理、分布存储和学习等神经网络的一些基本特征，从而确立了从系统的角度研究人工神经网络的基础.

1961 年 B. Windrow 和 M. E. Hoff 提出了自适应线性单元网络. 它可用于自适应滤波、预测和模型识别. 从上个世纪 50 年代到 60 年代初，神经网络的研究受到了人们的重视，研究工作进入一个高潮.

1969 年，美国麻省理工学院人工智能学者 M. Minsky 和 S. Papert 编写了影响很大的《Perception》一书. 该书指出：单层的感知机只能用于线性问题的求解，而对于像 XOR(异或)这样简单的非线性问题却无法求解. 他们还指出，能求解非线性问题的网络应该具有隐含层的多层神经网络，而将感知机模型扩展成多层网络是否有意义，还不能从理论上得到有力的证明. 由于 M. Minsky 的悲观结论，在这之后近 10 年中，神经网络的研究进入了一个缓慢发展的低潮期.

在这个低潮期，研究工作并没有完全停顿下来，仍有许多学者继续进行探索，并取得了一些重要成果. 在这期间，芬兰学者 T. Kohonen 提出了自组织映射理论；美国学者 S. A. Grossberg 提出了自适应共振理论(ART)；日本学者福岛邦彦(K. Fukushima)提出了认知机(Recognition)模型. 这些研究成果对以后神经网络的研究和发展都产生了重要影响.

美国加州理工学院生物物理学家 J. J. Hopfield 于 1982 年和 1984 年发表的两篇文章，有力地推动了神经网络的研究，引起了神经网络研究的又一次热潮. 1982 年他提出了一个新的神经网络模型——Hopfield 网络模型. 他在这种网络模型的研究中，首次引入了网络能量函数的概念，并给出了网络稳定性的判据. 1984 年 Hopfield 提出了网络模型实现的电子电路，为神经网络的工程实现指明了方向. Hopfield 的研究成果开拓了神经网络用于联想记忆和优化计算的新途径，并为神经计算机的研究奠定了基础. 1984 年 Hinton 等人将模拟退火算法引入到神经网络中，提出了 Boltzmann 机(BM)网络模型. BM 网络算法为神经网络优化跳出局部极小提供了一个有效的方法.

二、人工神经网络的研究现状

20 世纪 80 年代中期以来，神经网络的应用研究取得了很大的成绩，涉及的领域非常广泛. 就应用的技术领域而言有计算机视觉，语言的识别、理解与合成，优化计算，智能控制及复杂系统分析，模式识别，神经计算机的研制，知识处理，专家系统与人工智能. 涉及的学科有神经生理学，认识科学，数、理科学，心理学，信息科学，计算机科学，微电子学，光学，生物电子学等.

从众多应用研究领域取得的丰硕成果来看，人工神经网络的发展具有强大

的生命力.当前存在的问题是智能水平还不高,许多应用方面的要求还不能得到很好的满足;网络分析与综合的一些理论性问题(如稳定性、收敛性的分析,网络的结构综合等)还未得到很好的解决.随着人们对大脑信息处理机理认识的深化以及人工神经网络智能水平的提高,人工神经网络必将在科学技术领域发挥更大的作用.

针对当前人工神经网络存在的问题,今后研究的主要方向可分为理论研究和应用研究两个方面.在理论研究方面有:

(1) 利用神经生理与认知科学研究大脑思维及智能的机理.

(2) 利用神经科学基础理论的研究成果,用数理方法探索智能水平更高的人工神经网络模型,深入研究网络的算法和性能,例如稳定性、收敛件、容错性、鲁棒性等;开发新的网络数理理论,例如神经网络动力学、非线性神经场等.

在应用研究方面有:

(1) 神经网络软件模拟和硬件实现的研究.

(2) 神经网络在各个科学技术领域应用的研究.包括模式识别、信号处理、知识工程、专家系统、优化组合、智能控制等.

第二节　人工神经元模型

一、基本单元的神经元模型

图 5.1.1 表示了作为人工神经网络(artificial neural networks,以下简称ANN)的基本单元的神经元模型,它有三个基本要素:

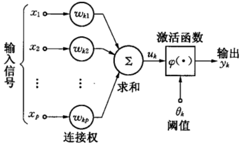


图 5.1.1　基本神经元模型

(1) 一组连接(对应于生物神经元的突触),连接强度由各连接上的权值表示,权值为正表示激活,为负表示抑制.

(2) 一个求和单元,用于求取各输入信号的加权和(线性组合).

(3) 一个非线性激活函数，起非线性映射作用并将神经元输出幅度限制在一定范围内(一般限制在(0, 1)或(−1, +1)之间).

此外还有一个阈值 θ_k(或偏值 $b_k = -\theta_k$)

以上作用可分别以数学式表达出来：

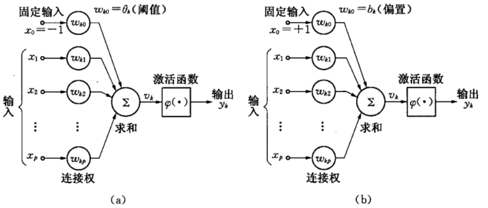
$$u_k = \sum_{j=1}^{p} w_{kj} x_j, \qquad v_k = net_k = u_k - \theta_k, \qquad y_k = \varphi(v_k),$$

式中 $x_1, x_2, \cdots, x_p$ 为输入信号，$w_{k1}, w_{k2}, \cdots, w_{kp}$ 为输入神经元 k 的信号的权值，u_k 为线性组合结果，θ_k 为阈值，$\varphi(\cdot)$为激活函数，y_k 为神经元 k 的输出.

若把输入的维数增加一维，则可把阈值 θ_k 也包括进去. 例如：

$$u_k = \sum_{j=0}^{p} w_{kj} x_j, \ y_k = \varphi(v_k).$$

此处增加一个新连接，其输入 $x_0 = -1$(或 $+1$)，权值 $w_{k0} = \theta_k$(或 b_k)，如图 5.1.2(a)、(b)所示.



(a) 包括阀值；(b) 包括偏值

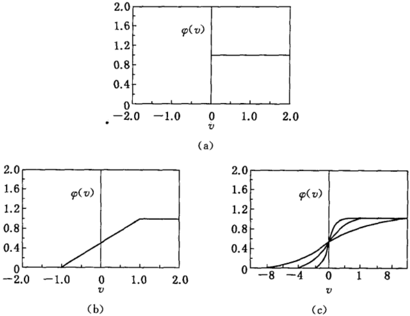
图 5.1.2 输入扩维后的神经元模型

二、激活函数

在神经元模型中的激活函数是对该神经元所有的输入信号求和之后的一个数学变换，一般激活函数可以有以下几种形式：

(1) 阈值函数(图 5.1.3(a))

$$\varphi(v) = \begin{cases} 1, & \text{当 } v \geqslant 0 \text{ 时}; \\ 0, & \text{当 } v < 0 \text{ 时}, \end{cases}$$



(a) 阈值函数;(b) 分段线性函数;(c) sigmoid 函数

图 5.1.3　激活函数

即阶梯函数. 这时相应的输出 y_k 为

$$y_k=\begin{cases}1, & \text{当 } v_k\geqslant 0 \text{ 时};\\ 0, & \text{当 } v_k<0 \text{ 时},\end{cases}$$

其中, $v_k=\sum_{j=1}^{p}w_{kj}x_j-\theta_k$. 常称此种神经元为 MP 模型.

(2) 分段线性函数(图 5.1.3(b))

$$\varphi(v)=\begin{cases}1, & \text{当 } v\geqslant 1 \text{ 时};\\ \dfrac{1}{2}(1+v), & \text{当 } -1<v<1 \text{ 时};\\ 0, & \text{当 } v\leqslant -1 \text{ 时}.\end{cases}$$

它类似一个放大系数为 1 的非线性放大器,当工作于线性区时它是一个线性组合器,放大系数趋于无穷大时变成一个阈值单元.

(3) sigmoid 函数(图 5.1.3(c))

最常用的函数形式为

$$\varphi(v)=\frac{1}{1+\exp(-\alpha v)},$$

参数 $\alpha > 0$ 可控制其斜率. 另一种常用的是双曲正切函数

$$\varphi(v) = \mathrm{th}\left(\frac{v}{2}\right) = \frac{1-\exp(-v)}{1+\exp(-v)},$$

这类函数具有平滑的渐近线,并保持单调性.

第三节 人工神经网络的表示

如果把上一节的神经元模型用信号流图(signal flow graph)表示,则可更清楚地显示出其作用过程.

信号流图是由一些带方向的连接和结点组成的,信号沿连接线按箭头方向流动,有两种连接线:突触连接(图 5.1.4(a))和函数连接(图 5.1.4(b)). 结点也有两种:求和结点(图 5.1.4(c))和分送结点(图 5.1.4(d)).

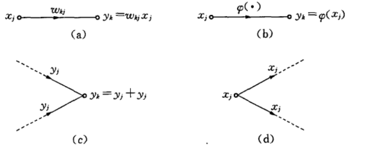


图 5.1.4 信号流图的单元

因此也可把 ANN 理解为一个由带有突触连接和函数连接的结点组成的有向图,图 5.1.5 表示了神经元内部的信号流动. 当只研究各神经元之间的信号流动时,可把内部的信号流图省去(图 5.1.6),用同样方式也可以表示神经网络.

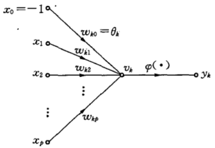


图 5.1.5 神经元内部的信号流图

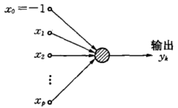


图 5.1.6 神经元作为基本单元

第四节　网络结构及工作方式

在神经网络中除单元特性外，网络的拓扑结构也是一个重要特性. 从连接方式看 ANN 主要有两种类型：

(1) 前馈型网络. 各神经元接受前一层的输入，并输出给下一层，没有反馈（图 5.1.7）. 结点分为两类，即输入单元和计算单元，每一计算单元可有任意个输入，但只有一个输出（它可耦合到任意多个其他结点作为输入）. 通常前馈网络可分为不同的层，第 i 层的输入只与第 $i-1$ 层输出相连，输入和输出结点与外界相连，而其他中间层称为隐层.

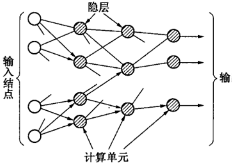


图 5.1.7　前馈网络

(2) 反馈型网络. 所有结点都是计算单元，同时也可接受输入，并向外界输出，可画成一个无向图（图 5.1.8(a)），其中每个连接弧都是双向的，也可画成图 5.1.8(b)的形式. 若总单元数为 n，则每一个结点有 $n-1$ 个输入和一个输出.

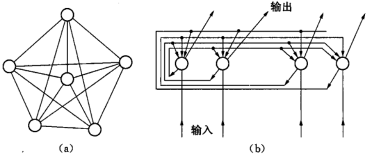


图 5.1.8　单层全连接反馈网络

ANN 的工作过程主要分为两个阶段：第一阶段是学习期，此时各计算单元状态不变，各连接线上的权值可通过学习来修改；第二阶段是工作期，此时各连

接权固定，计算单元变化，以达到某种稳定状态.

从作用效果看，前馈网络主要是函数映射，可用于模式识别和函数逼近. 反馈网络按对能量函数的极小点的利用来分类，有两类：第一类是能量函数的所有极小点都起作用，这一类主要用作各种情况的联想存储器；第二类只利用全局极小点，它主要用于求解最优化问题.

第五节　人工神经网络的学习

一、学习方式

通过向环境学习获取知识并改进自身性能是 ANN 的一个重要特点，在一般情况下，性能的改善是按某种预定的度量调节自身参数（如权值）随时间逐步达到的，学习方式（按环境所提供信息的多少分）有以下三种.

(1) 监督学习（有教师学习）

这种学习方式需要外界存在一个"教师"，可对一组给定输入提供应有的输出结果（正确答案）. 这组已知的输入—输出数据称为训练样本集. 学习系统可根据已知输出与实际输出之间的差值（误差信号）来调节系统参数（图 5.1.9）.

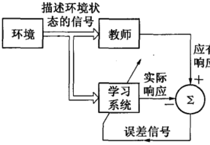


图 5.1.9　监督学习框图

(2) 非监督学习（无教师学习）

非监督学习时不存在外部教师，学习系统完全按照环境所提供数据的某些统计规律来调节自身参数或结构（这是一种自组织过程），以表示外部输入的某种固有特性（如聚类，或某种统计上的分布特征）（图 5.1.10）.

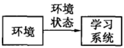


图 5.1.10　非监督学习框图

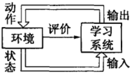


图 5.1.11　再励学习框图

(3) 再励学习(强化学习)

这种学习介于上述两种情况之间,外部环境对系统输出结果只给出评价(或奖罚)而不是给出正确答案,学习系统通过强化那些受奖励的动作来改善自身性能(图 5.1.11).

二、学习算法

(1) 误差纠正学习

令 $y_k(n)$为输入 $x_j(n)$时神经元 k 在 n 时刻的实际输出,$d_k(n)$表示应有的输出(可由训练样本给出),则误差信号可写为

$$e_k(n) = d_k(n) - y_k(n).$$

误差纠正学习的最终目的是使某一基于 $e_k(n)$的目标函数达到最小,以使网络中每一输出单元的实际输出在某种统计意义上最逼近于应有输出. 一旦选定了目标函数形式,则误差纠正学习就成为一个典型的最优化问题. 最常用的目标函数是均方误差判据,定义为

$$J = \mathbf{E}\left(\frac{1}{2}\sum_k e_k^2(n)\right),$$

其中 $\mathbf{E}$ 是求期望算子. 上式的前提是被学习的过程是平稳的,具体方法可用最速梯度下降法. 直接作为目标函数时,需要知道整个过程的统计特性,为解决这一困难,可用误差在时刻 n 的瞬时值$\varepsilon(n)$代替 J,即

$$\varepsilon(n) = \frac{1}{2}\sum_k e_k^2(n).$$

则问题变为求 $\varepsilon(n)$对权值 w 的极小值,据最速梯度下降法可得

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n),$$

其中 η 为学习步长. 这就是通常说的误差纠正学习规则(或称 delta 规则).

(2) Hebb 学习

神经心理学家 Hebb 提出的学习规则可归结为"当某一突触(连接)两端的神经元的激活同步(同为激活或同为抑制)时,该连接的强度应增加,反之则应减弱",用数学方式可描述为

$$\Delta w_{kj}(n) = F(y_k(n),\ x_j(n)),$$

式中 $y_k(n)$, $x_j(n)$分别为 w_{kj} 连接两端神经元的状态,其中最常用的一种情况为

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n).$$

由于 $\Delta w_{kj}(n)$ 与 $y_k(n)$，$x_j(n)$ 相关成比例，有时称之为相关学习规则.

(3) 竞争学习

顾名思义，在竞争学习时网络各输出单元互相竞争，最后达到只有一个最强者激活. 最常见的一种情况是输出神经元之间有侧向抑制性连接(图 5.1.12)，这样众多输出单元中如有某一单元较强，则它将获胜并抑制其他单元，最后只有比较强者处于激活状态. 最常用的竞争学习规则可写为

$$\Delta w_{kj}(n) = \begin{cases} \eta(x_j - w_{ji}), & \text{若神经元竞争获胜；} \\ 0, & \text{若神经元竞争失败.} \end{cases}$$

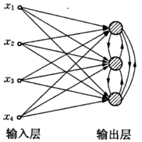


图 5.1.12 具有侧向连接的竞争学习网络

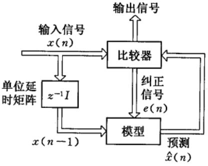


图 5.1.13 自适应系统框图

三、学习与自适应

当学习系统所处环境平稳时(统计特征不随时间变化)，从理论上说通过监督学习可以学到环境的统计特征，这些统计特征可被学习系统(神经网络)作为经验记住. 如果环境是非平稳的(统计特征随时间变化)，通常的监督学习没有能力跟踪这种变化，为解决此问题需要网络有一定的自适应能力，此时每一个不同输入都作为一个新的例子对待，其工作过程如图 5.1.13. 此时模型(如 ANN)被当作一个预测器，基于前一时刻输出 $x(n-1)$ 和模型在 $n-1$ 时刻的参数，它估计出 n 时刻的输出 $\hat{x}(n)$，$\hat{x}(n)$ 与实际值 $x(n)$(作为应有的正确答案)比较，其差值 $e(n)$ 称为“新息”，如新息 $e(n)=0$，则不修正模型参数，否则应修正模型参数以便跟踪环境的变化.

第二章　人工神经网络基本模型及算法

人工神经网络基本模型主要是指早期几个有代表性的模型. 本章将介绍 MP 模型、感知器模型、BP 算法和 Hopfield 网络模型.

第一节　MP　模　型

MP 模型属于一种阈值元件模型，它是由美国 McCulloch 和 Pitts 提出的最早神经元模型之一，MP 模型是大多数神经网络模型的基础.

一、标准 MP 模型

通常考虑某一神经元要受到其他神经元的作用，因而总是以 n 个神经元相互连接形成神经元计算模型(图 5.2.1). 一个神经元具备相应的输入和输出，但是神经元自身的状态决定其输出的有无. 即是，每一个神经元从其他 $n-1$ 个神经元接受信息，产生神经兴奋和冲动，在其他条件不变的情况下，不论何种刺激，只要达到阈值以上，就能产生一个动作电位，并以最快速度作非衰减的等幅传递输出；一旦输入的总和小于阈值，神经元处于抑制状态，没有被激活，也就没有任何输出产生.

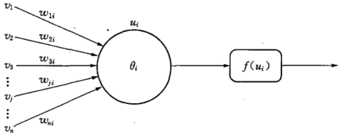


图 5.2.1　MP 神经元模型

对 n 个互连的神经元中的第 i 个神经元，外界输入的总和影响其激活值. i 神经元的状态以某种函数形成输出，即有

$$u_i = \sum_{j=1}^{n} w_{ji} v_j - \theta_i ,$$

$$v_i = f(u_i), \tag{1}$$

其中，w_{ji}代表神经元i与神经元j之间的连接强度(模拟生物神经元之间突触连接强度)，称之为连接权；u_i代表神经元i的活跃值，即神经元i的状态；v_i代表神经元i的输出；θ_i代表神经元i的阈值.

函数f表达了神经元的输入输出特性. 在MP模型中，f定义为阶跃函数

$$v_i = \begin{cases} 1, & \text{当 } u_i > 0 \text{ 时}; \\ 0, & \text{当 } u_i \leqslant 0 \text{ 时}. \end{cases}$$

如果把阈值θ_i看作为一个特殊的权值，则式(1)可改写为

$$v_i = f\left(\sum_{j=0}^{n} w_{ji} v_j\right),$$

其中，$w_{0i} = -\theta_i$，$v_0 = 1$.

为用连续型的函数表示神经元的非线性变换能力，常采用S型函数

$$f(u_i) = \frac{1}{1 + \mathrm{e}^{-u_i}},$$

该函数的图像如图5.1.3(c)所示.

MP模型在发表时并没有给出一个学习算法来调整神经元之间的连接权. 但是，我们可以根据需要，采用一些常见的算法来调整神经元之间的连接权，以达到学习目的. 上一章介绍的Hebb学习规则就是一个较好的学习算法.

二、延时MP模型

延时MP模型可表示为

$$v_i(t) = f\left(\sum_{j=1}^{n} w_{ji} v_j(t - \tau_{ji}) - \theta_i\right),$$

其中，τ_{ji}为突触时延. 该模型具有下面的一些特点：

(1) 神经元的状态满足"激活/抑制"规律，即0/1律；

(2) 神经元为一"多输入/单输出"处理单元；

(3) 具有"空间整合"与"阈值"作用，即

$$v_i(t) = \begin{cases} 1, & \text{当 } \sum\limits_{j \neq i} w_{ji} v_j(t - \tau_{ji}) > \theta_i \text{ 时}; \\ 0, & \text{当 } \sum\limits_{j \neq i} w_{ji} v_j(t - \tau_{ji}) \leqslant \theta_i \text{ 时}; \end{cases}$$

(4) 所有神经元具有相同的、恒定的工作节律，工作节律取决于突触时延τ_{ji}；

(5) 没有考虑时间整合作用和不应期；

(6) 神经元突触的时延 τ_{ji} 为常数，权系数也为常数，即

$$w_{ji} = \begin{cases} 1, & \text{当 } v_j \text{ 为兴奋性输入时；} \\ -1, & \text{当 } v_j \text{ 为抑制性输入时.} \end{cases}$$

三、改进的 MP 模型

对于上面的 MP 模型，如考虑时间整合作用和不应期，则有

$$v_i(t) = f\left(\sum_{i=1}^{n} w_{ii} v_i(t - k\tau_{ii}) + \sum_{j \neq i}^{n} w_{ji} v_j(t - k\tau_{ji}) - \theta_i\right), k = 1, 2, \cdots$$

其中，$w_{ji}(k)$ 随 k 变化（也即随时延变化），且

$$w_{ji}(k) \stackrel{\text{def}}{=} \begin{cases} > 0, & \text{若兴奋性突触；} \\ \leqslant 0, & \text{若抑制性突触.} \end{cases}$$

这里，$\sum_{j \neq i}^{n} w_{ji} v_j(t - k\tau_{ji})$ $(k = 1, 2, \cdots)$ 表示对过去所有输入进行时间整合，而 $w_{ji}(k)$ 表示神经元内的反馈连接权.

$$w_{ii}(k) = \begin{cases} -\alpha, & \text{当 } \theta_i = \infty \text{ 时} \quad \text{（绝对不应期）；} \\ -h(k), & \text{当 } \beta < \theta_i < \infty \text{ 时} \quad \text{（相对不应期）；} \\ 0, & \text{当 } \theta_i \leqslant \beta \text{ 时} \quad \text{（反应期）.} \end{cases}$$

式中 α 为正数，$h(k)$ 为 k 的单调减的指数函数. 改进的 MP 模型与延时 MP 模型的主要区别在于其结构的可塑性，即在改进的 MP 模型中，权系数是可变的，譬如兴奋性突触的连接权将增大或减小.

第二节　感知器模型

感知器是一种早期的神经网络模型，由美国学者 F. Rosenblatt 于 1957 年提出. 由于在感知器中第一次引入了学习的概念，使人脑所具备的学习功能在基于符号处理的数学模型中得到了一定程度的模拟，所以引起了广泛的关注.

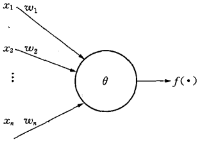


图 5.2.2　感知器处理单元

一、简单感知器

简单感知器模型（图 5.2.2）实际上仍然是 MP 模型的结构，但是它通过采用监督学

习来逐步增强模式划分的能力，达到所谓学习的目的.

感知器处理单元对 n 个输入进行加权和操作，即

$$y = f(\sum_{i=1}^{n} w_i x_i - \theta), \tag{2}$$

其中，w_i 为第 i 个输入到处理单元的连接权值，θ 为阈值.

作为分类器，感知器的输出 $y = f(\cdot)$ 通常取阶跃函数：

$$y = \begin{cases} 1, & 当 \sum_{i=1}^{n} w_i x_i - \theta > 0 时; \\ 0, & 当 \sum_{i=1}^{n} w_i x_i - \theta \leqslant 0 时. \end{cases} \tag{3}$$

感知器在形式上与 MP 模型差不多，它们之间的区别在于神经元之间连接权的变化. 感知器的连接权定义为可变的，这样感知器就被赋予了学习的特性.

利用简单感知器可以实现逻辑代数中的一些运算，如表 5.2.1 所示.

表 5.2.1　简单逻辑运算

x_1	x_2	$y = x_1 \wedge x_2$	$y = x_1 \vee x_2$	$\bar{x}_1$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0

对应简单感知器的计算式为

$$y = f(w_1 x_1 + w_2 x_2 - \theta). \tag{4}$$

(1) “与”运算，即 $x_1 \wedge x_2$.

当取 $w_1 = w_2 = 1, \theta = 1.5$ 时，式(4)完成逻辑“与”运算.

(2) “或”运算，即 $x_1 \vee x_2$.

当取 $w_1 = w_2 = 1, \theta = 0.5$ 时，式(4)完成逻辑“或”运算.

(3) “非”运算，即 $\bar{x}_1$.

当取 $w_1 = -1, w_2 = 0, \theta = -1$ 时，式(4)完成逻辑“非”运算.

和许多代数方程一样，式(3)中不等式具有一定的几何意义. 对于一个两输入的简单感知器，每个输入取值为 0 和 1，如上面给出的逻辑运算，所有样本输入有四个，记为 (x_1, x_2)：(0, 0)，(0, 1)，(1, 0)和(1, 1)，其构成样本输入空间.

例如，在二维平面上，对于"或"运算，各个样本的分布如图 5.2.3 所示. 直线 $x_1 + x_2 - 0.5 = 0$ 将二维平面分为两部分，上部为激发区（$y = 1$，用★表示），下部为抑制区（$y = 0$，用☆表示）.

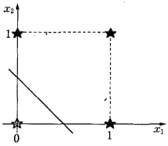


图 5.2.3　逻辑"或"的二维表示

感知器的训练（学习）是通过监督学习过程来实现的. 样本矢量 $\boldsymbol{X}$ 输入到感知器中，感知器中的各处理单元按式(2)或式(3)计算，得到实际输出 y，将其与期望输出 d 的差来修正下一步计算的权值. 简单感知器引入的学习算法称之为误差学习算法. 该算法是神经网络学习中的一个重要算法，并已被广泛应用，该算法的计算步骤为：

(1) 选择一组初始权值 $w_i(0)$.

(2) 计算某一输入模式对应的实际输出与期望输出的误差 δ.

(3) 如果小于给定误差限则结束，否则继续下一步.

(4) 更新权值（阈值可视为输入恒为 1 的一个权值）

$$\Delta w_i(t+1) = w_i(t+1) - w_i(t) = \eta[d - y(t)]x_i,$$

式中 η 为在区间(0, 1)上的一个常数，称为学习步长，它的取值与训练速度和收敛的稳定性有关；d, y 为神经元的期望输出和实际输出；x_i 为神经元的第 i 个输入.

(5) 返回(2)，重复，直到对所有训练样本模式，网络输出均能满足要求.

但简单感知器无法解决"异域"问题，所以简单感知器只能进行线性分类，而对非线性样本的划分无法实现. 感知器对线性不可分问题的局限性决定了它只有较差的归纳性，而且通常需要较长的离线学习才能达到收敛. 然而，感知器的线性分类性质决定了它能存贮丰富的线性可分的输入输出模式对，另外感知器的优点还表现在它对模式的直接联想和容易理解的工作方式. 要解决线性不可分模式的分类问题，需要对网络进一步改进.

感知器所提出的分类算法在神经网络研究中有着重要的意义和地位，它具有神经网络一些自组织、自学习的特征. 对于线性可分的模式空间，其算法是严格收敛的. 在后来的神经网络研究中，许多模型都是在这种指导思想下建立的，或者是这种模型的改进和推广.

二、多层感知器

如果在输入和输出层之间加上一层或多层的神经元（隐层神经元），就可构成多层前向网络，叫做多层感知器（图 5.2.4）.

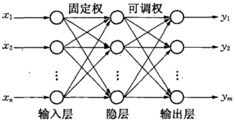


图 5.2.4　多层感知器

这里需指出的是：多层感知器只允许调节一层的连接权．这是因为按感知器的概念，无法给出一个有效的多层感知器学习算法．图5.2.4的三层感知器中，有两层连接权，输入层与隐层单元间的权值是随机设置的固定值，不被调节；输出层与隐层间的连接权是可调节的．

对于上一章述及的 XOR 问题，用一个简单的三层感知器就可得到解决，解决方法如图 5.2.5 所示．

$$x_1^{(1)} = f(1 \cdot x_1^{(0)} + 1 \cdot x_2^{(0)} - 0.5),$$

$$x_2^{(1)} = f((-1) \cdot x_1^{(0)} + (-1) \cdot x_2^{(0)} - (-1.5)),$$

$$y = f(1 \cdot x_1^{(1)} + 1 \cdot x_2^{(1)} - 1.5).$$

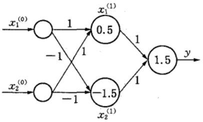


图 5.2.5　XOR 求解网络

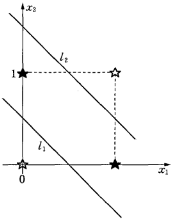


图 5.2.6　XOR 的空间划分

实际上，该三层感知器的输入层和隐层的连接，就是在模式空间中用两个超平面去划分样本（图 5.2.6），即为两条直线

$$l_1: x_1 + x_2 = 0.5, \qquad l_2: x_1 + x_2 = 1.5.$$

输出层相当于一个逻辑"与"运算，这样就构成了图 5.2.6 所示的两条直线

间的区域，从而完成这种非线性模式划分. 可以证明，只要隐层和隐层单元数足够多，多层感知器网络可实现任何模式分类. 但是，多层网络的权值如何确定，即网络如何进行学习，在感知器上没有得到解决. 当年 Minsky 等人就是因为对于非线性空间的多层感知器学习算法未能得到解决，使其对神经网络的研究作出悲观的结论.

第三节 BP 算法

一、多层前向 BP 网络

由于在神经网络中引入隐层神经元，神经网络就具有更好的分类和记忆能力，因此相应的学习算法就成了研究的焦点. 1985 年 Rumelhart 等提出的 EBP (Error Back Propagation)算法(简称 BP)，系统地解决了多层神经元网络中隐单元层连接权的学习问题，并在数学上给出了完整的推导.

采用BP 算法的多层神经网络模型一般称为BP 网络(图 5.2.7). 它由输入层、中间层和输出层组成，中间层也就是隐含层，可以是一层或多层结构.

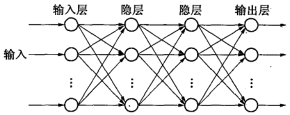


图 5.2.7 多层前向 BP 网络

BP 网络的学习过程由两部分组成：正向传播和反向传播. 当正向传播时，输入信息从输入层经隐单元层处理后传向输出层，每一层神经元的状态只影响下一层的神经元状态. 如果在输出层得不到希望的输出，则转入反向传播，将误差信号沿原来的神经元连接通路返回. 返回过程中，逐一修改各层神经元连接的权值. 这种过程不断迭代，最后使得信号误差达到允许的范围之内.

BP 网络中采用了有一定阈值特性的连续可微的 sigmoid 函数作为神经元的激发函数. 当然其他类似的非线性函数也可选用. 这里采用 S 型(Sigmoid)函数

$$f(x) = \frac{1}{1+e^{-x}}.$$

二、sigmoid 激发函数下的 BP 算法

由于 BP 网络要求采用非线性的连续可导的激发函数，经多年试验研究，现

通常采用S型函数作为激发函数

$$f(Net_{kj})=\frac{1}{1+e^{-Net_{kj}}},$$

式中，Net_{kj}为网络单元u_j的状态

$$Net_{kj}=\sum_i w_{ji}o_{ki}+\theta_j.$$

则单元输出为

$$o_{kj}=\frac{1}{1+\exp(-\sum_i w_{ji}o_{ki}-\theta_j)}=\frac{1}{1+e^{-Net_{kj}}},$$

其中，θ_j为单元u_j的阈值.在这种激发函数下，有

$$f_j'(Net_{kj})=\frac{\partial o_{kj}}{\partial Net_{kj}}=o_{kj}(1-o_{kj}).$$

故对输出层单元有训练误差

$$\delta_{kj}=o_{kj}(1-o_{kj})(t_{kj}-o_{kj}),$$

对隐层单元有训练误差

$$\delta_{kj}=o_{kj}(1-o_{kj})\sum_m \delta_{km}w_{mj},$$

权值调节为

$$\Delta w_{ji}(t+1)=\eta\delta_{kj}o_{ki}.$$

在实际的学习过程中，学习速率η对学习过程的影响很大.η是按梯度搜索的步长.η越大，权值的变化越剧烈.实际应用中，通常是以不导致振荡的前提下取尽量大的值.为了使学习速度足够快而不易产生振荡，往往在规则中再加上一个“势态项”，即

$$\Delta w_{ji}(t+1)=\eta\delta_{kj}o_{ki}+\alpha\Delta w_{ji}(t),$$

式中，α是一个常数，它决定过去权重的变化对目前权值变化的影响程度.

下面给出算法步骤：

(1) 置各权值或阈值的初始值：$w_{ji}(0)$，$\theta_j(0)$为小的随机数值.

(2) 提供训练样本，输入矢量$\boldsymbol{X}_k=(x_{k_1},x_{k_2},\cdots,x_{k_M})$，$k=1,2,\cdots,p$，$M$为输入向量的维数；期望输出$\boldsymbol{T}_k(t_{k_1},t_{k_2},\cdots,t_{k_N})$，$k=1,2,\cdots,p$，$N$为输出向量的维数；对每个输入样本进行下面(3)到(5)的迭代.

(3) 计算网络的实际输出及隐层单元的状态

$$o_{kj} = f_j(w_{ji}o_{ki} + \theta_j).$$

(4) 计算训练误差

$$\delta_{kj} = o_{kj}(1 - o_{kj})(t_{kj} - o_{kj}) \quad (输出层);$$

$$\delta_{kj} = o_{kj}(1 - o_{kj})\sum_m \delta_{km}w_{mj} \quad (隐含层).$$

(5) 修正权值和阈值

$$w_{ji}(t+1) = w_{ji}(t) + \eta\delta_j o_{ki} + \alpha[w_{ji}(t) - w_{ji}(t-1)],$$

$$\theta_j(t+1) = \theta_j(t) + \eta\delta_j + \alpha[\theta_j(t) - \theta_j(t-1)].$$

(6) 当 k 每经历 1 至 p 后，判断指标是否满足精度要求：

$$E \leqslant \varepsilon,$$

其中，$E = \frac{1}{2}\sum_k \sum_j (t_{k_j} - o_{k_j})^2$，$\varepsilon$ 为精度要求. 若满足则训练结束，否则继续进行(3)到(5)的迭代训练，直到满足精度要求为止.

BP 模型虽然在各方面都具有重要意义，而且应用也很广泛，但它也存在一些不足. 从数学上看，它是一个非线性优化问题，不可避免地存在局部极小点；学习算法的收敛速度很慢；网络隐层结点数选取带有很大的盲目性和经验性，尚无理论上的指导；新加入的样本要影响已学完的样本等.

第四节　HP 网络模型

一、反馈网络

在反馈网络中，所有结点都是一样的，它们之间都可以互相连接(一个结点既接受其他结点来的输入，同时也输出给其他结点). 图 5.2.8(a)画出这种网络的展开图，图 5.2.8(b)是简化形式(为简单计，在 5.2.8(a)中认为神经元是 MP 模型，图中所示为各结点没有自反馈的情况). 由于引入反馈，所以它是一个非线性动力学系统.

对于一个非线性动力学系统，系统的状态从某一初值出发经过演变后可能有如下几种结果：

(1) 渐近稳定点(吸引点)；

(2) 极限环；

(3) 混沌(chaos)；

(4) 状态发散.

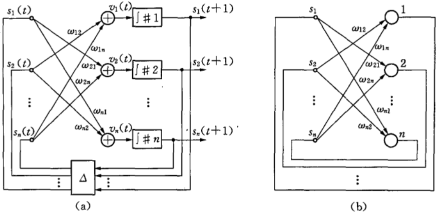


图 5.2.8 反馈网络

由于神经元模型是一个有界函数，故系统的状态不会发生发散现象. 本节不讨论理论分析这类系统的方法，而从应用角度对最常用的反馈网络——Hopfield 网络模型给出一些定性分析.

二、离散 Hopfield 网络

离散的 Hopfield 网络构造如图 5.2.8 所示，作用过程可写为

$$\begin{cases} v_j(t) = \sum_{i=1}^{N} w_{ji}s_i(t) - \theta_j, \\ s_j(t+1) = \operatorname{sgn}[v_j(t)], \end{cases} \quad \text{即} \quad s_j(t+1) = \begin{cases} +1, & \text{当 } v_j(t) > 0 \text{ 时;} \\ -1, & \text{当 } v_j(t) < 0 \text{ 时.} \end{cases}$$

一般认为 $v_j(t) = 0$ 时神经元状态保持不变，N 为网络结点总数.

将上式合并可写为

$$s_j(t+1) = \operatorname{sgn}\left(\sum_{i=1}^{N} w_{ji}s_j(t) - \theta_j\right), \tag{5}$$

其中 $s_j(t)$为任一时刻单元 j 的状态(取+1 或—1). 整个网络的状态可用列向量 $\boldsymbol{s}$ 表示

$$\boldsymbol{s} = (s_1, s_2, \cdots, s_N)^{\mathrm{T}}.$$

一般情况下网络是对称的($w_{ji} = w_{ij}$)，所以 $\boldsymbol{W}$ 可用一个 $N\times N$ 的对角线为 0 的对称矩阵表示. 这种网络有两种工作方式

(1) 串行(或称异步，asynchronous)方式：任一时刻只有一个单元按式(5)改变状态，其余单元保持不变(各单元动作顺序可以随机选择，或按某种确定顺序动作).

(2) 并行(或称同步,synchronous)方式:某一时刻所有神经元同时改变状态.

如果网络从 $t=0$ 的任一初态 $s(0)$ 开始变化,存在某一有限 t,从此以后网络状态不再变化,即

$$s(t+1)=s(t),$$

则称网络式(5)达到稳定状态.显然稳定状态应满足

$$s_j=\operatorname{sgn}\left(\sum_{i=1}^{N}w_{ji}s_i-\theta_j\right)\quad(j=1,2,\cdots,N).$$

定义网络的能量函数为

$$E=-\frac{1}{2}\sum_i\sum_j w_{ji}s_i(t)s_j(t)-\sum_i s_i(t)\theta_i\quad\text{(异步方式)},$$

$$E=-\frac{1}{2}\sum_i\sum_j w_{ji}s_i(t+1)s_j(t)-\frac{1}{2}\sum_i[s_i(t)+s_i(t+1)]\theta_i\quad\text{(同步方式)},$$

$$\text{则 }|E|\leqslant\frac{1}{2}\sum_i\sum_j|w_{ji}|\cdot|s_i|\cdot|s_j|+\sum_i|s_i|\cdot|\theta_i|$$

$$=\frac{1}{2}\sum_{ij}|w_{ji}|+\sum_i|\theta_i|.$$

由于 s_i, s_j 只可能是 $+1$ 和 -1, w_{ij} 和 θ_i 均有界,所以能量函数也是有界的,即从任一初始状态开始,若在每次迭代时都满足 $\triangle E=E(t+1)-E(t)\leqslant 0$,则网络的能量将越来越小,最后趋向于稳定状态 $\triangle E=0$.

三、连续 Hopfield 网络

连续的 Hopfield 网络可与一电子线路对应(图 5.2.9),每一神经元可用一个(有正反向输出的)放大器模拟,输入端并联的电阻和电容可模拟生物神经元的时间常数,互相连接间的电导 T_{ij} 则模拟各神经元间突触的特性(相当于权系数).

该网络的微分方程为

$$\begin{cases}C_i\dfrac{\mathrm{d}v_i}{\mathrm{d}t}=\displaystyle\sum_{i=1}^{N}T_{ji}s_i-\frac{v_j}{R_j}+i_j,\\ s_j=g(v_j).\end{cases}$$

可证明,若 g^{-1} 为单调增且连续, $C_j>0$, $T_{ji}=T_{ij}$,则沿系统的轨迹有 $\dfrac{\mathrm{d}E}{\mathrm{d}t}\leqslant 0$;当且仅当 $\dfrac{\mathrm{d}v_i}{\mathrm{d}t}=0$ 时, $\dfrac{\mathrm{d}E}{\mathrm{d}t}=0$, $i=1,2,\cdots,N$. 其中

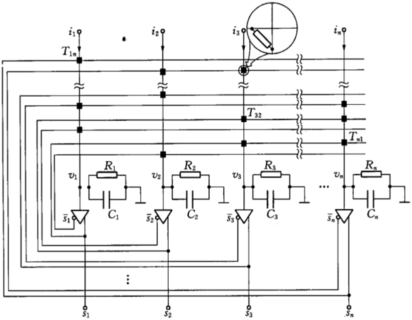


图 5.2.9　连续 Hopfield 网络对应的电路图

$$E=-\frac{1}{2}\sum_i\sum_j T_{ij}s_is_j-\sum_j i_js_j+\sum_j\frac{1}{R}\int_0^{s_j}g_j^{-1}(v_j)\mathrm{d}v_j$$

为系统的能量函数.

可以证明,在一定条件下,能量函数 E 是一个单调减函数.所以当从某一初始状态变化时,网络的演变是使 E 下降,达到某一局部极小时就停止变化.这些能量的局部极小点就是网络的稳定点或称吸引子.

第三章　人工神经网络设计与应用

本章在前两章介绍人工神经网络基本理论、模型和算法的基础上，对当前人工神经网络在汉字识别、信号处理和自动控制中的应用进行介绍，具有一定的代表性和实践指导意义.

第一节　人工神经网络在手写体汉字识别中的应用

一、引言

近年来人们提出了很多实现汉字识别的方法和技术，由于汉字在手写体形式中的特点，现有手写体汉字识别系统需要进一步改进以提高其效率，它们缺乏足够的灵活性和并行处理能力. 特别是在光照条件差、噪声污染严重的情况下系统就变得不够实用和有效. 重新兴起的神经网络和并行计算技术给人们以启示：神经网络可实现传统(精确)逻辑推理所不能完成的模糊语义逻辑推理；相应的并行计算结构可提高信息处理速度，从而使手写体汉字识别系统成为办公自动化中的实用工具. 为此，本节介绍一种神经网络算法以实现手写体汉字识别的逻辑分析功能.

二、算法介绍

该算法将二值化图像的非监督、非线性特征提取与贝叶斯统计决策的非监督分类、反馈学习机制相结合. 在传统方法中，特征提取和目标识别是两个独立的过程，它们之间的联系未得到充分的利用. 但从人类认知心理学角度来看，人脑的生物神经网络系统是以集成方式工作的，视觉信息输入的特征被视网膜神经元阵列所感知，并被并行分布式结构映射为时变参数集. 本节有别于传统算法，它以自组织、自学习和自适应机制模拟人类联想功能和认知映射，相应认知(识别决策)结果由神经网络算法输出，其中自组织分类算法由 Kohonen 网络构成，该分类算法可处理原始图像阵列，并包括以联想记忆、分布式存贮方式实现的特征隐式提取. 而 Kohonen 网络的知识学习与权值更新是由多层 BP 网络实现的，这也可认为是对 Kohonen(传统)网络的一种改进. 该算法的结构框图如图 5.3.1 所示.

自组织神经网络可自动地通过外界环境进行学习，它是与人类感知过程一

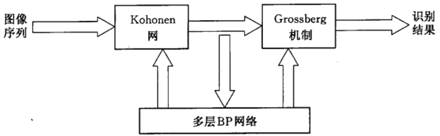


图 5.3.1　算法结构框图

致的.该神经网络的非监督学习特性可被认为是来源于其自组织机制，而像 BP 这样的神经网络是有教师监督的学习能力很强的自适应系统，它适合于样本学习和特定类识别，该算法是通过 BP 学习算法来增强其非线性能力的.

三、手写体汉字识别算法

1. 神经网络算法介绍

根据图 5.3.1 所示的结构，神经网络算法描述如下.

(1) 输入层和输出层

原始像素阵列被用作神经网络的输入，该层的每一个结点被组织成多维向量形式.模式集合被映射成超球面上的点集.识别结果由输出层给出，它具有贝叶斯意义下的最佳值.

(2) 中间层(Kohonen 层)结点.

Kohonen 网络输出结点 $\overline{\boldsymbol{Y}}$ 的量化反映了输入向量 $\overline{\boldsymbol{X}}$ 的匹配程度.

最佳输出 $\overline{Y}_{opt}(\in \{\overline{Y}_i(t) \mid i=0,\ 1,\ \cdots,\ N\})$ 应满足输入模式和权向量之间的最优匹配.根据该规则，权向量调整如下

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{W}_{ij}(t)) = \alpha(t) * (\boldsymbol{X}_i(t) - \boldsymbol{W}_{ij}(t)), \qquad i \in N_C(t),$$

其离散形式为

$$\boldsymbol{W}_{ij}(t+1) = \boldsymbol{W}_{ij}(t) + \alpha(t) * (\boldsymbol{X}_i(t) - \boldsymbol{W}_{ij}(t)), \qquad i \in N_{Ej},\ 0 \leqslant i \leqslant 1.$$

其中 $\alpha(t)$ 为增益系数，N_C 和 N_{Ej} 分别表示公式中结点相应的领域.

该网络所应用的 Kohonen 结构得到 20 世纪 70 年代以来视觉感知方面生物神经网络研究的实验支持.

(3) BP 层结点

在多层 BP 网结构中，映射被用以调整和训练 Kohonen 结点集(权值集)：

$$F: \mathbf{R}^n \to \mathbf{R}^n,\ \overrightarrow{\boldsymbol{W}}_{BP} = F(\overline{\boldsymbol{Y}}),$$

其中$\vec{W}_{BP}$为$\vec{W}_{\text{Kohonen}}$矩阵的最佳逼近.

(4)"引入层"

Grossberg 学习原理引导 Kohonen 结点对最终识别的决策过程，Kohonen 结点的最终权值调整是由"引入层"的宏结点完成的：

$$W_{\text{Kohonen}}^{new} = W_{\text{Kohonen}}^{old} + (d * W_{BP}^{new} - e * Y),$$

其中 e, d 分别表示小于 1 的常数，并随训练时间 t 的增加而减小.

2. 识别过程

预处理过程包括图像输入、汉字分割、标准化($n \times n$ 阵列，n 由用户选定)和骨架化(Skeletonizing，即二值化过程)，然后表征原始汉字信息的像素阵列被用作该算法的输入. 由于每个汉字的点阵尺寸可固定为若干常数，所以该算法的通常规格为 16×16、32×32 等(该规格是由标准化过程所决定的). 识别过程的第一步是对随机得到的初步识别结果进行迭代推理，直至获得最佳值集合为止.

传统方法是从图像阵列中提取特征向量，但是通常情况下，特征向量并不能有效、准确和充分地表达汉字属性，特别是在手写体汉字识别系统中，断笔和连笔就是实用系统中需要解决的比较困难的问题. 到目前为止，像句法模式识别这样的基于模型的方法仍不能完善地解决该问题. 其根本原因就是传统人工智能技术缺乏容错能力和有关模式信息不完全情况下的模糊推理功能. 而人工神经网络恰恰在该方面具有很强的优势. 从本质上讲，Kohonen 结构实现了模式信息的自适应压缩，并使汉字骨架特征的非形式化表示得到增强，以并行分布化的网络结点形式记忆其本质属性. 同时，BP 网络通过非线性测度空间映射和逼近，减少了模式分类系统中冗余的信息. 由 BP 网和 Kohonen 网综合而构成的联想记忆表征机制是本节中所提出的新算法与现有算法的一个重要区别.

3. 系统的硬件支持与软件仿真环境

该神经网络的实现需要具有并行分布式处理的硬件支持，可用于软件环境仿真的系统体系结构可提供功能强大的仿真外壳(Shell)和并行运算功能，它是使系统实用化的重要环节.

第二节　人工神经网络在通讯信号处理中的应用

一、信号检测

在通讯时，接收到的信号会受到噪声影响. 在带噪声的接收信号中如何分辨

是哪一种信号或者有没有所需的信号并非易事，特别是当噪声分布是任意时. 设在时间区间$(-T/2,\ T/2)$内，

$$y(t) = s_i(t) + n(t),\ i = 0,\ 1;\ |t| \leqslant T/2,$$

其中，$y(t)$是接收到的信号，$s_i(t)$是两种可能的原始信号，$n(t)$是加性噪声. 假设对接收信号进行均匀采样，采样间隔的长度保证了离散的噪声是白噪声. 此外还假设系统的频带比信号的频带宽得多. 在时间段$(-T/2,\ T/2)$的末端，接受器必须以观测的信号$y(t)$的样本作出决策，判断所收到的信号到底是$s_0(t)$还是$s_1(t)$.

这样的接收器可以用图 5.3.2 所示的二输入无记忆装置来构造.

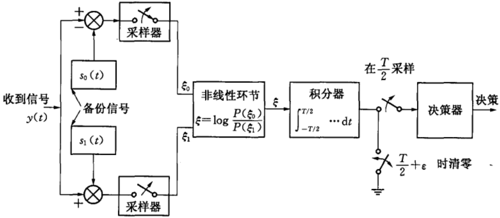


图 5.3.2 信号接收器

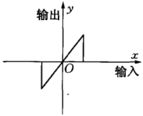


图5.3.3 “穿孔”函数曲线

对于正态分布白噪声，最佳接收器是线性装置. 对于指数型或 Laplace 型分布的噪声，分段线性也就足够了. 如图 5.3.3 中实线所示的“穿孔”函数就是其中典型的一种.

把有可能传输的信号备份起来，分别与收到的信号相减，在采样时刻 t_k，有

$$\xi_i(t_k) = y(t_k) - s_i(t_k), \qquad i = 0,\ 1.$$

这些值送至二输入装置，得到的输出是

$$\xi(t_k) = \ln \frac{p(\xi_0(t_k))}{p(\xi_1(t_k))}.$$

在$(-T/2,\ T/2)$上进行积分，在时刻 $T/2$ 与一个阈值相比较，作出最终决策. 当信号 $s_i(t)$ $(i = 0,\ 1)$ 的水平与噪声比足够小时，可证明最优接收器的输

入特性是

$$f(x)=\frac{\mathrm{d}}{\mathrm{d}x}\ln(p(x)).$$

信号 $s_i(t)(i=0,1)$ 是两极形连续或分段连续信号时，最优接收器可解释为一个无记忆非线性环节后接一个匹配滤波器. 如果 $p(x)$ 是高斯分布，$f(x)$ 是线性函数，接收器也就是典型的线性接收器，可以采用自适应线性合成器(ALC).

对于图 5.3.3 实线所示的非线性"穿孔"，用 $1\times20\times5\times1$ 结构的网络，用 BP 算法训练 50 000 次就能达到很满意的效果.

对于其他噪声分布，可采用多层前传网络来构成最优接收器(图 5.3.4).

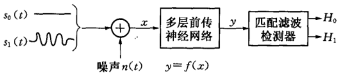


图 5.3.4 神经网络最优接收器

图 5.3.4 中的两个输入信号 $s_0(t)$ 实际上是无信号，而另一个信号 $s_1(t)$ 是一段固定长度的正弦型信号，频率、幅度、相位都是已知的.

在水声和极低频通讯中常常会有脉冲噪声的出现，放电、点火火花可能是引起这种脉冲的原因. 大气噪声、某些人为噪声、潜艇上的发电机的噪声、潜艇靠近海岸时陆地上的电站对潜艇通讯的干扰都可能是脉冲型的. 在通讯中，这类噪声有可能传播数千千米之远，因此是一种重要类型的噪声.

具有这种噪声的接收信号 $y(t)$，通过由神经网络构成的非线性环节，再通过匹配滤波器决定接受对应于无输入的假设 H_0 还是正弦信号的假设 H_1，匹配滤波器把非线性输出与已知时间校准信号相乘再积分，其值大于阈值就选 H_1，否则就是 H_0.

为了仿真，设噪声是零均值 Laplace 分布的方差 $\sigma=14.1$ 的脉冲噪声，两脉冲间隔 ΔT 服从 Poisson 分布，其均值依所加噪声信噪比不同而不同. 当信噪比从 1 dB 到 −24 dB 时，其值为 500 至 2 个采样间隔不等. 信号 $s_1(t)$ 的持续时间是 250 个样本、幅度为 1.0 的正弦波.

如果噪声的概率分布密度 $p(x)$ 的解析表达式已知，当然也就不需要用到神经网络，直接构成一个 $f(x)$ 公式的非线性环节就可以了. 即使一定要用神经网络构造，也只需将已知的 $f(x)$ 的输入输出数据用于训练，构造出图 5.1.3 所示的非线性函数即可. 但一般的情况是，噪声的概率分布甚至噪声的类型对设计者来说是未知的. 为此采用图 5.3.5 所示的自适应训练形式.

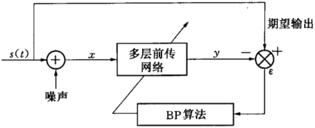


图 5.3.5　网络自适应训练框图

图 5.3.5 中的多层感知机仍采用 $1\times20\times5\times1$ 结构，初始权重取小的随机数. BP 算法中学习步长 $\eta=0.1$，动量系数 $\alpha=0.5$. 为了比较，当噪声为零均值 Gauss 分布时($\sigma=0.1$)，训练出的多层感知机在信号幅度±1 的范围内呈现出线性关系，在训练 3 000～5 000 次后这种关系即很明显. 当噪声是上述脉冲噪声，而信噪比 SNR =－5dB，$\Delta T=125$ 时，在训练 5 000 次后就开始形成较明显的非线性环节，对超过原信号电平的带噪声信号进行剪波. 当训练次数逐渐增至 10 000 次时，非线性函数向“穿孔”函数逼近.

这样的信号检测系统的效果如何呢? 用训练 10 000 次后的结果网络来进行测试. 检测信号的信噪比为5 dB～－24 dB的 2 000 个信号. 仿真表明这样的检测出错率几乎接近于零，只有当 SNR 超过－24 dB 时出错率才开始迅速增加. 如果我们在噪声分布一无所知的情况下采用线性环节处理，那么当信噪比从 5 dB降至－24 dB 时出错率将从零增至 36%. 可见，这种方法是十分有效的.

从所形成的非线性来看，当接收信号的幅度超过原始信号的振幅±1 的范围时，接收器进行了抑制，因为大的输入可能是由于脉冲噪声引起的，它所包含的信息量低，模糊度高，因此对它抑制有利于最终判决.

二、信道均衡

在一些通讯设备如电话、无线电装置、光纤通讯中，信号通道上的频带范围内的频率响应常常不够平坦，相位响应也可能是非线性的. 当数字信号高速通过这些通道时，由于传播媒介的扩散性，常常会产生所谓的“符号间干涉”现象，降低传输的准确性.

如果在通讯时首先对输入信号滤波，具体地说，就是建立一个自适应滤彼器，它本身是通道的逆系统，那么，串起来后就可以补偿通道的幅度和相位的不规则性，提高传输准确率. 为了解决这一问题，可以引入由神经网络构成的自适应均衡器.

自适应均衡器由多个时延环节及其抽头出来的自适应线性组合器件(ALC)组成，加权之和得出反卷积信号，量化成二进制输出. 这样就与传过通道

的原始二进制数据在形式上一致起来. ALC 和量化器件一起组成了一个 Adaline 神经元. 本网络的权重可以用任何形式的最小二乘类算法来调节，其中 LMS 算法是最著名的一种，目前几乎所有的长途通讯中都采用了这种算法. 在均衡器中具体实现这种算法是由 AT&T 实验室的 Lucky 发明的，这种方法被称为"面向决策的学习". 如果没有均衡器，那么通道二进制量化输出所产生的数据流的传输误差小于 10^{-1}，因此，可以把它作为期望响应来训练神经元. 在训练的开始阶段，由于期望响应中的噪声作用，有时可能会使得适应过程朝着错误的方向进行，但这只是零星现象，就平均趋势来讲，适应过程是正确进行的. 当神经元学到一定程度后，期望响应中的噪声也就消除了. 这个自适应均衡器收敛后，这种出错率一般不超过 10^{-6}. 这就保证了在同样的可靠性条件下，采用自适应均衡器能够以高于没有均衡器的 4 倍速度传输数据.

在长途电话网中，还有一个需要解决的问题是回波现象. 产生这种现象的原因是在线路的两个方向上都有放大作用，而且线路各端的传话机和接收机是串联耦合的. 为了解决这个问题，一开始是在会议广播系统单通道通讯中采用了回波抑制器，隔断了反馈作用. 为了使声音和数据能够同时双通道传送，目前已广泛采用自适应回波抵消器.

在解决了上述两个技术问题后，目前国际上的长途通讯中正在发展和应用集成服务数字网络. 这种网络在每个终端都使用了自适应均衡器和自适应回波抵消器. 从而可以在普通的铜质电话线路上进行高速通讯.

第三节　人工神经网络在自动控制中的应用

一、神经网络模型参考自适应控制

控制的基本目的就是对一给定的物理过程提供合适的输入，使该过程产生所期望的输出. 这个物理过程称之为对象，而构成合适对象输入的装置称为控制器. 图 5.3.6 是一般控制系统方框图. 从图 5.3.6 中看出，前馈控制器直接根据期望输出产生信号，而反馈控制器根据期望输出与实际对象之差来产生控制信号.

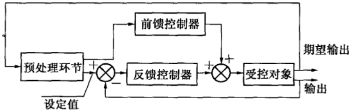


图 5.3.6　一般控制系统框图

控制系统的稳定性是最基本的要求，因此，反馈是控制的内核. 在控制对象与环境愈来愈复杂的情况下，我们希望控制器能够根据对象行为的观测量，自适应地控制对象至期望要求. 神经网络自适应控制器在这方面显示了独特的优势. 它通过不断地学习，获取对象的知识并适应过程的变化. 神经控制器的设计也像传统的控制器一样有两种方式：一是通过系统辨识获取对象的数学模型，再根据一定的设计指标即控制要求设计；另一条途径是根据对象的输出误差调节控制器内部参数以直接达到控制的目的. 在控制理论中，这两种思想分别称为间接控制和直接控制. 神经网络控制中，分别对应于泛化学习与特定学习.

泛化学习实际上就是利用对象的输入输出数据学习对象的逆系统或预报模型. 在训练阶段随意地产生控制器的输出信号，提供给对象，由此信号及对象的真实输出来训练控制器，最终使得控制器能产生正确的控制信号，使对象的输出尽可能地接近期望的轨迹. 由于对象是动态的因果系统，要想神经控制器完全等同于它的逆系统是困难的. 这种控制器的训练与控制阶段显然是分开的，训练阶段不能保证对象的输出不超过正常的工作范围. 另外，由于它的立论是假设对象的动态性质固定不变，因此，它还不是自适应的.

特定学习与此不同，它将控制器的设计直接与控制精度挂钩，利用对象的真实输出与期望输出之差来调整神经网络控制器的权重. 这是一个边控制边学习的过程，因此，具有适应对象与环境变化的能力.

为了使泛化学习具有自适应能力，可将模型参考自适应控制（MRAC）与之结合，由此构成所谓神经网络 MRAC 控制（图 5.3.7）. 图 5.3.7 中，TDL 表示将当前时刻信号进行若干步延时. 神经网络 N^i 是对非线性对象的辨识，其目的是

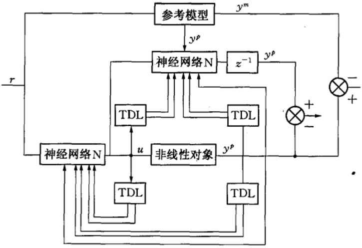


图 5.3.7　神经网络 MRAC 控制框图

利用当前及以前时刻对象的输入输出数据来预报下一步对象的输出应该是什么，预报准确程度用 $e^i(k+1)=y^p(k+1)-y(k+1)$ 来衡量. 由于预报模型精确度不够及反馈可能引入的系统失稳，还要在控制过程中依训练准则 J 进一步辨识

$$J=\sum_{k=0}^{T_i}\|e^i(t+k)\|^2.$$

作为控制器参数的设计，是在假设这个辨识的模型精确地描述了对象的特征的情况下进行的，这就是间接控制这一词的来历.

MRAC 的任务是确定控制信号 $\{u(k)\}$，使得相同参考输入下对象的输出 $y(k)$ 与参考模型的输出 $y^m(k)$（即期望输出 $y_k(k)$）的差不超过给定的误差范围. 用公式表示为

$$\lim_{k\to\infty}\|y(k)-y^m(k)\|\leqslant\varepsilon,$$

其中，ε 是一个给定的小正数.

引入神经网络后，如果辨识的模型中控制 $u(k)$ 是显式的非线性映射，即 $u(k)$ 可显式地表示成 $y^p(k+1)$, $y(k)$, $\cdots$, $y(k-n)$, $u(k-1)$, $\cdots$, $u(k-n)$，则可直接将辨识模型构成控制器. 如果 $u(k)$ 与 $y^p(k+1)$, $y(k)$, $\cdots$, $y(k-n)$, $u(k-1)$, $\cdots$, $u(k-n)$ 不能方便地表达出来，还应该再引入一个神经网络 Nc 进行训练，直至上述关系完全明了. 控制网络 Nc 的训练应根据控制效果进行，训练数据为

$$J^c=\sum_{k=0}^{T^c}\|e^c(t+k)\|^2, \tag{1}$$

即网络参数经过 T^c 步后调整一次. T^c 和 T^i 的选择对于整个方案很重要. 根据传统的自适应控制理论和经验，取 $T^c>T^i$，即控制作用频繁些，而相对地辨识则"以静制动"，更能把握对象的规律. 如果 T^c 和 T^i 都取较大的值，则可能因为对对象的"不闻不问"而失去对对象的控制，输出误差最终发散.

例 设被控系统的真实模型为

$$y(k+1)=f(y(k),\ y(k-1))+g(u(k)),$$

其中

$$f(y(k),\ y(k-1))=\frac{y(k)y(k-1)(y(k)+2.5)}{1+y^2(k)+y^2(k-1)}.$$

参考模型为

$$y^m(k+1)=0.6y^m(k)+0.2y^m(k-1)+r(k),$$

其中，$r(k)$ 是参考输入信号.

当 $g(u(k))=u(k)$ 时，选用 $2\times 20\times 10\times 1$ 结构的网络 $N(\cdot)$ 辨识 $f(\cdot)$，辨识模型为

$$\hat{y}^p(k+1)=N(y(k),\ y(k-1))+u(k).$$

为了保证系统是输入输出有界的，辨识用输入信号选取为 $(-2,\ 2)$ 内均匀分布的白噪声序列.

若采用如下控制规律

$$u(k)=-N(y(k),\ y(k-1))+0.6y(k)+0.2y(k-1)+r(k),$$

则控制误差

$$e^c(k+1)=0.6e^c(k)+0.2e^c(k-1).$$

显然系统是渐近稳定的，故 $\lim\limits_{k\to\infty}e^c(k)=0$.

当 $g(u(k))=u^3(k)$，而且函数 $f(\cdot)$ 和 $g(\cdot)$ 对设计都未知时，问题要复杂些. 辨识模型取为

$$\hat{y}^p(k+1)=N^f(y(k),\ y(k-1))+N^g(u(k)),$$

控制规律为

$$u(k)=-\hat{g}^{-1}\{-N^f(y(k),\ y(k-1))+0.6y(k)+r(k)\}.$$

设 N^f 的结构是 $2\times 20\times 10\times 1$，$N^g$ 的结构是 $1\times 20\times 10\times 1$，都已经训练出来了，但因 N^g 复杂，不能直接写出 g^{-1}. 现取 $u(k)$ 为 $(-4,\ 4)$ 内均匀分布的白噪声，以使得和 N^g 在广泛的非线性范围内作用. 训练神经网络控制器 N^c，使得 $N^g(N^c(r))=r$，从而

$$u(k)=N^c\{-N^f(y(k),\ y(k-1))+0.6y(k)+r(k)\}.$$

参考输入可取

$$r(k)=\sin\left(2\pi\frac{k}{25}\right)+\sin\left(2\pi\frac{k}{10}\right).$$

网络的训练可以采用通常的 BP 算法，也可以采用动态 BP 算法.

在非线性自适应控制中，一个未解决的难题就是当控制系统的逆系统不存在时的处理问题. 以一阶系统为例，设

$$y(k+1)=f(y(k))+g(u(k)),$$

如果 $g(u)$ 为某一值时，同时有多个自变量 u 存在，那么 g^{-1} 是不存在的. 例如

$$g(u)=(u+1)u(u-1),$$

当 $u=-1$，$u=0$ 和 $u=1$ 时都有 $g(u)=0$，那么当 $g(u)=0$ 时，u 的取值到

底是上述 3 种中的哪一个呢？若采用动态 BP 算法，难题有望得到解决.

首先，我们假设能够获取足够精确的网络 $N^f(y(k))$ 和 $N^g(u(k))$，使得模型 $\hat{y}(k+1) = N^f(\hat{y}(k)) + N^g(u(k))$ 以足够的精度逼近对象模型. 使用多层网络控制器 N^c 产生控制信号，它的训练不是根据它与逆系统的逼近程度，而是直接根据辨识模型的输出 $y(k)$（注意不是对象的输出）与参考模型的输出 $y^m(k)$ 之差 $\hat{e}^c(k) = \hat{y}(k) - y^m(k)$. 控制网络的训练准则是式(1). 这种控制结构见图 5.3.8.

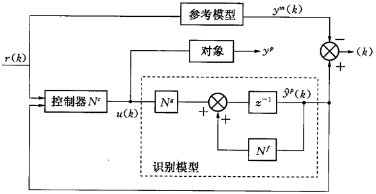


图 5.3.8 不可逆系统 MRAC 控制器结构

如果 $g(u)$ 的逆本来就存在，这种方法与前面的逆系统方法同样能达到控制要求. 当逆系统不存在时，就须采用动态 BP 算法.

例如，系统模型

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + (u(k)+1)u(k)(u(k)-1),$$

参考模型

$$y^m(k+1) = 0.6y^m(k) + r(k),$$

参考输入

$$r(k) = a\left(\sin\left(2\pi\frac{k}{25}\right) + \sin\left(2\pi\frac{k}{10}\right)\right).$$

可取 $a = 0.25$，N^c 的结构是 $2\times20\times10\times1$，$T^c = 1$，$\eta = 0.001$，训练 100 000 步后模型输出 $y^m(k)$ 与网络输出 $y(k)$ 达到一致. 若 a 的取值增大，则系统的非线性作用范围更大，应增加训练时间.

以上研究表明，神经网络间接自适应控制器的设计是成功的. 训练时间长是 BP 算法时固有的缺点，实际应用中，可以用一定长度的样本反复训练，以减少在线训练时间.

二、神经网络自校正控制

前面讲了神经网络的模型参考自适应控制，仿真表明它们在处理非线性系统的模型不确定性等方面有显著优点. 我们知道，在自适应控制中，还有一类重要的方法就是自校正控制(STC)，下面就看看如何将 STC 与神经网络结合起来.

考虑如下一类可反馈线性化的单输入单输出系统

$$y(k+1)=f(y(k),\cdots,y(k-n))+g(y(k),\cdots,y(k-n),u(k-1),\cdots,u(k-n))u(k)$$

其中，$\{u(k)\}$和$\{y(k)\}$分别是系统的输入、输出，$f(\cdot)$和 $g(\cdot)$是非线性函数，假定非线性函数$g(\cdot)$的逆存在，即它在感兴趣的范围内不为零.

如果函数 $f(\cdot)$和 $g(\cdot)$是非线性函数，那么，可轻易地采用下面的控制，使得输出达到我们所期望的轨迹 $y_d(k+1)$，即采用

$$u(k)=\frac{-f(y(k),\cdots,y(k-n),u(k-1),\cdots,u(k-n))+y_d(k+1)}{g(y(k),\cdots,y(k-n),u(k-1),\cdots,u(k-n))}.$$

当 $f(\cdot)$和 $g(\cdot)$未知时，可以采用两个神经网络，通过学习来逼近. 各网络的输入输出神经元是线性的，而隐含层神经元特性是 S 形曲线. 逼近$f(\cdot)$的神经网络权重 w^f，其第一层权重记为 w_i^f，第二层权重记为 v_i^f，$i=0,1,2,\cdots,q^f$，q^f 为隐含神经元个数；逼近$g(\cdot)$的神经网络权重 w^g，其第一层权重记为 w_i^g，第二层权重记为 v_i^g，$i=0,1,2,\cdots,q^f$.

为简单起见，考虑如下一阶系统

$$y(k+1)=f(y(k))+g(y(k))u(k).$$

当引入神经网络后，$f(y(k))$用 $\hat{f}(y(k),w^f(k))$代替. 系统的输出预报值应为

$$\hat{y}(k+1)=\hat{f}(y(k),w^f(k))+\hat{g}(y(k),w^g(k))u(k).$$

为使系统的输出接近期望的输出，取

$$\hat{y}(k+1)=y_d(k+1),$$

这时，加入对象的控制信号采用

$$u(k)=\frac{y_d(k+1)-\hat{f}(y(k),w^f(k))}{\hat{g}(y(k),w^g(k))}.$$

在这样的输入下，记期望输出 $y_d(k+1)$ 与对象真实输出之差为

$$e(k+1)=y(k+1)-y_d(k+1).$$

如果逼近 $f(\cdot)$和 $g(\cdot)$的两个神经网络完全精确地描述了$f(\cdot)$和

$g(\cdot)$,那么,在没有系统噪声干扰的情况下,输出误差 $e(k+1)$ 应恒为零. 因此,可采用

$$J = e^2(k+1)$$

作为训练两个网络的判据. 调节 w^f 和 w^g,使得 J 的值减小. 可以推导出

$$\frac{\partial J}{\partial w^f(k)} = \frac{g(y(k))}{\hat{g}(y(k),\ w^g(k))} \cdot \frac{\partial \hat{f}(y(k),\ w^f(k))}{\partial w^f(k)} e(k+1)$$

及

$$\frac{\partial J}{\partial v^g(k)} = -\frac{g(y(k))}{\hat{g}(y(k),\ w^g(k))} \cdot \frac{\partial \hat{g}(y(k),\ w^g(k))}{\partial w^g(k)} u(k)e(k+1),$$

其中,$\frac{\partial \hat{f}}{\partial w^f}$、$\frac{\partial \hat{g}}{\partial w^g}$分别是 $\hat{f}$ 对 w^f 的各个标量权重及 $\hat{g}$ 对 w^g 的各个标量权重求偏导组成的

$$\frac{\partial \hat{f}(y(k),\ w^f(k))}{\partial w_i^f(k)} = \sigma(w_i^f(k)y(k)), \quad i = 0,\ 1,\ 2,\ \cdots,\ q^f,$$

$$\frac{\partial \hat{f}(y(k),\ w^f(k))}{\partial w_i^f(k)} = \sigma'(w_i^f(k)y(k))v_i^f(k)y(k), \quad i = 0,\ 1,\ 2,\ \cdots,\ q^f,$$

而

$$\frac{\partial \hat{g}(y(k),\ w^g(k))}{\partial w_i^g(k)} = \sigma(w_i^g(k)y(k)), \quad i = 0,\ 1,\ 2,\ \cdots,\ q^g,$$

$$\frac{\partial \hat{g}(y(k),\ w^g(k))}{\partial w_i^g(k)} = \sigma'(w_i^g(k)y(k))v_i^g(k)y(k), \quad i = 0,\ 1,\ 2,\ \cdots,\ q^g,$$

其中,$\sigma(\cdot)$是隐含神经元特性,为 S 形曲线.

上面只是针对一阶系统这种简单情形作了初步的分析,这种方法可以进一步推广到更广泛的系统中去,原理与上面的方法是大致一样的.

三、神经网络自适应控制

在自适应控制理论的发展历史上共走了两条路,一条是间接自适应控制器的设计,一条是直接自适应控制器的设计. 对于第一条路,就线性系统来讲,在过去的 20 年内已取得了很大的进步,而第二条路却没有走通.

何谓直接自适应控制呢? 直接自适应控制就是直接依据对象的知识来调整控制器的内部参数,以使得对象的输出误差尽量小,它与间接自适应控制不同之处就在于不需要预先或同时辨识对象的模型. 引入神经网络后,我们可以进行这一工作.

图 5.3.9 是神经网络直接自适应控制器框图.

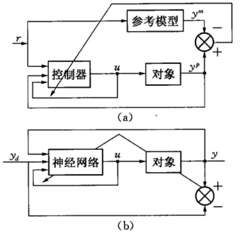


图 5.3.9 神经网络直接自适应控制器框图

图 5.3.9(a)是基于 MRAC 的,图 5.3.9(b)则是以 STC 为基础的.神经网络可以是多层前传网络,也可以是小脑模型关节控制器(CMAC)神经网络模型等其他网络.这里仅考虑图 5.3.9(b),采用多层前传网络和 BP 算法训练.

非线性控制系统的描述为

$$y(k+1)=f(y(k),\ u(k)).$$

如果允许自由增加 $y(k)$的维数,上述表达式总是成立的.

作为控制器的神经网络,它的训练准则应取为

$$J_1=\|u(k)-u_d(k)\|^2,$$

$$J_2=\sum_{k=1}^{\infty}\|u(k)-u_d(k)\|^2,$$

其中,$u(k)$为神经网络控制器的输出量,亦即对象的控制输入,对应的对象输出为 $y(k)$, $u_d(k)$是当期望的对象输出为$y_d(k)$时,神经网络应提供的控制量.

作为对控制的要求,希望对象的输出达到给定的轨迹$y_d(k)$.怎样才能知道这时对象应具有的输入$u_d(k)$的值呢? 在对象未知的情况下做不到这一点.而$u_d(k)$又是控制用神经网络训练必须的,需要避开这个矛盾.

我们可以把神经网络控制器与对象看作一个整体,即一个具有更多层数的神经网络.网络的最后一层(或几层)是固定的,用它来描述对象,当然也是未知的.这个整体神经网络的训练准则改为以输出误差为基准,即取

$$J = \| y(k+1) - y_d(k+1) \|_Q^2 = \| e^c(k+1) \|_Q^2.$$

现在的目的很明确，选取当前控制量 $u(k)$，使下一步(或多步)的输出误差尽可能小. 上式中 Q 是对 $y(k+1) - y_d(k+1) = e^c(k+1)$ 的各分量的加权矩阵，为对称正定阵，为简单起见，取为时间 k 的常数阵.

设神经网络控制器某个可修改的参数 $\theta \in \Theta$, Θ 为参数空间，θ 为标量，根据梯度法调整各个 θ 的值.

为了计算$\frac{\partial J}{\partial \theta(k)}$，当 $\theta(k)$ 是网络最后一层权重时，采用如下规则

$$\frac{\partial J}{\partial \theta(k)} = \frac{\partial J}{\partial u^{\mathrm{T}}(k)} \cdot \frac{\partial u(k)}{\partial \theta(k)}.$$

而对应于 J 的梯度$\frac{\partial J}{\partial u^{\mathrm{T}}(k)}$的值为

$$\frac{\partial J}{\partial u^{\mathrm{T}}(k)} = \frac{\partial J}{\partial y^{\mathrm{T}}(k+1)} \cdot \frac{\partial y(k+1)}{\partial u^{\mathrm{T}}(k)} = 2(e^c(k+1))^{\mathrm{T}} Q \cdot \frac{\partial y(k+1)}{\partial u^{\mathrm{T}}(k)}.$$

在上述计算中，包含了对象的 Jacobi 矩阵 $\frac{\partial y(k+1)}{\partial u^{\mathrm{T}}(k)}$. 而对于未知的对象来说，它是未知的. 但是，如果通过先验知识能确定出$u(k)$各分量加入后 $y(k+1)$ 的各分量的变化趋势，即能对$u(k)$影响 $y(k+1)$ 的方向有所了解，那么就有办法避开这个问题. 记$u(k)$各分量对 $y(k+1)$ 各分量影响的符号组成的矩阵为 $\mathrm{sgn}\left(\frac{\partial y(k+1)}{\partial u^{\mathrm{T}}(k)}\right)$，记

$$\frac{\partial J_s}{\partial u^{\mathrm{T}}(k)} = 2(e^c(k+1))^{\mathrm{T}} Q \cdot \mathrm{sgn}\left(\frac{\partial y(k+1)}{\partial u^{\mathrm{T}}(k)}\right).$$

如果网络收敛，只要 $y(k+1) \neq y_d(k+1)$，就有 $\frac{\partial J_s}{\partial u^{\mathrm{T}}(k)} \neq 0$.

对于网络的最后一层，$\theta = w_{ij}^{L-1}$，即参数 θ 为最后一层的权重，则

$$\frac{\partial u_j(k)}{\partial \theta(k)} = \sigma'(u_j(k)) y_j^{L-1}(k),$$

其中，$u_j(k)$为向量 $u(k)$的第 j 个分量，即神经网络最后一层各输出节点的输出值；$y_i^{L-1}(k)$ 为第 $L-1$ 层隐含节点的输出值；$\sigma(\cdot)$是隐含神经元特性，为 S 形曲线.

根据以上分析知，网络最后一层的权重修改

$$w_{ij}^{L-1}(k+1) = w_{ij}^{L-1}(k) - \eta \delta_i^L(k) y_i^{L-1}(k),$$

其中 $$\delta_j^L(k) = 2\sigma'(u_j(k))(e^c(k+1))^T Q \cdot \operatorname{sgn}\left(\frac{\partial y(k+1)}{u_j(k)}\right).$$

对于其他各层的权重调整，仍按标准的 BP 算法进行.

由上面方法得出一种新的自适应控制，有如下两个显著特点：第一，没有特定的学习阶段，即没有依赖于被控系统的数学模型的辨识，而用了少量定性的先验知识，控制器是直接根据控制效果设计的；第二，控制器的参数调整是一个依时间关系的自适应过程. 因此，这种控制器属于直接自适应控制器.

值得一提的是，已有的仿真实验表明，即使 Jacobi 矩阵真的是已知的，并不需要用符号代替，用算法训练时在收敛速度上也并无明显的优越之处.

参考文献

1. 王洪元，史国栋. 人工神经网络技术及其应用. 北京：中国石化出版社，2002
2. 杨建刚. 人工神经网络实用教程. 杭州：浙江大学出版社，2001
3. 胡守仁，余少波，戴葵. 神经网络导论. 长沙：国防科技大学出版社，1993
4. 胡守仁，余少波，戴葵. 神经网络应用技术. 长沙：国防科技大学出版社，1993
5. 刘健勤. 手写体汉字识别的神经网络算法. 计算机应用研究，1994，27(3)：6-7

第六篇　小波变换

小波变换是数学发展史上的重要成果,它无论对数学还是对工程应用都产生了深远的影响,小波变换已经广泛应用于理论数学、应用数学、信号处理、语音识别与合成、自动控制、图像处理与分析等领域. 从原则上讲,凡传统使用 Fourier 变换的地方,都可用小波变换来代替,小波变换在时域频域同时具有良好的局部化特性,而且由于对高频采取逐渐精细的时域或空域步长,从而可以聚焦到分析对象的任意细节,对传统的 Fourier 变换提出了挑战.

本篇详细论述了小波变换的原理和基本方法,阐述其在信号检测、语音信号处理等领域的应用.

第一章　Fourier 变换

本章介绍了在 Fourier 级数中周期函数可由离散频率为 $n\Delta\omega$ 的谐波来表现,接下来介绍了非周期函数可由谐波来表现的连续 Fourier 变换,而为了用计算机作频谱分析,需要把信号离散量和频谱函数离散量作为 Fourier 变换对,建立对离散量的相应算法,这就是第三节介绍的离散 Fourier 变换,在本章的最后一节介绍了具有时—频局部化功能的窗口 Fourier 变换.

第一节　Fourier 级数

一、Fourier 级数

众所周知,以 T 为周期的任意连续信号 $f(t) = f(t+T)$ 可被视为某种复杂的波,物理实验表明,它可由形如 $A_n\sin(nt+\theta)$ 的若干谐波叠加而成,换句话说,周期性的模拟信号 $f(t)$ 可分解为不同频率、振幅和相位的谐波信号. 因此,$f(t)$ 可表示为如下形式

$$\begin{aligned} f(t) &= \sum_{n=0}^{\infty} A_n \sin(nt+\theta) \\ &= \sum_{n=0}^{\infty} (A_n \sin\theta \cos nt + A_n \cos\theta \sin nt) \\ &= \sum_{n=0}^{\infty} (a_n \cos nt + b_n \sin nt), \end{aligned}$$

其中，$A_n \sin(nt+\theta)$ 表示 n 次谐波，其振幅是 A_n，n 表示频率，θ 表示相位. 经过积分以及变量代换后，可得到连续的周期函数 $f(t) = f(t+T)$ 的 Fourier 级数形式

$$\begin{cases} f(t) = \dfrac{a_0}{2} + \sum\limits_{n=1}^{\infty} (a_n \cos n\Delta\omega t + b_n \sin n\Delta\omega t), \\ a_n = \dfrac{2}{T}\displaystyle\int_{-T/2}^{T/2} f(t) \cos n\Delta\omega t \, \mathrm{d}t, n = 0, 1, \cdots \\ b_n = \dfrac{2}{T}\displaystyle\int_{-T/2}^{T/2} f(t) \sin n\Delta\omega t \, \mathrm{d}t, n = 1, 2, \cdots \\ \Delta\omega = 2\pi/T. \end{cases} \tag{1}$$

对式(1)作简单形变，$f(t)$的 Fourier 级数还可表现为另一种形式，即

$$f(t) = \sum_{n=0}^{\infty} A_n \sin(n\Delta\omega t + \theta_n), \tag{2}$$

其中，$A_n = \sqrt{a_n^2 + b_n^2}$ 和 $\theta_n = \arctan(b_n/a_n)$ 分别表示了频率为$n\Delta\omega/(2\pi)$的谐波振幅和相位.

式(1)与式(2)表明，周期为 T 的函数确实可分解为若干谐波之和，但这些谐波的频率是离散出现的，不是连续出现的，并且频率是基频$\Delta\omega$的整数倍.

二、可用 Fourier 级数表示的函数

函数(模拟信号)可能有间断，这类函数可以展开为 Fourier 级数吗?

定理　设 $f(t)$是以 T 为周期的实函数，且在$[-T/2, T/2]$上连续或仅有有限个第一类间断点，允许有限个极值点而不允许无穷振荡，则在$f(t)$的连续点处，$f(t)$可由式(1)表示；在$f(t)$的间断点处，$f(t) = [f(t-0)+f(t+0)]/2$，即取左、右端值的中间值.

若仅在$[-T/2, T/2]$上有一段信号 $f(t)$，可从两个方面理解它的 Fourier 级数. 一方面是关于$f(t)$表现为级数的，可理解$f(t)$作周期延拓成周期为 T 的函数，由定理知，$f(t)$可表示为 Fourier 级数；另一方面是关于 Fourier 级数表现实际信号$f(t)$的，虽然级数在 **R** 上都有表现，但在$[-T/2, T/2]$区段上才是表

现真实信号的. 另外,虽然在间断点处出现中间值,实际信号并无此表现,但这不会影响对原信号的讨论.

对一段信号 $f(t)$, $t \in [0, T/2]$ 延拓成周期为 T 的函数,可作奇延拓,也可作偶延拓. 若作奇延拓,则$f(t)$是周期为 T 的函数,在$[-T/2, T/2]$上是奇函数,它的 Fourier 级数中仅含$\sin n\Delta\omega t$项;若作偶延拓,$f(t)$在$[-T/2, T/2]$上是偶函数,它的 Fourier 级数中仅含$\cos n\Delta\omega t$项.

三、Fourier 级数的复指数形式

利用复变函数的基本知识,将

$$\cos n\Delta\omega t = \frac{1}{2}(\mathrm{e}^{-\mathrm{i}n\Delta\omega t} + \mathrm{e}^{\mathrm{i}n\Delta\omega t}),$$

$$\sin n\Delta\omega t = \frac{\mathrm{i}}{2}(\mathrm{e}^{-\mathrm{i}n\Delta\omega t} - \mathrm{e}^{\mathrm{i}n\Delta\omega t})$$

代入式(1),就有

$$\begin{cases} f(t) = \sum\limits_{n=-\infty}^{+\infty} C_n \mathrm{e}^{\mathrm{i}n\Delta\omega t}, \\ C_n = \dfrac{1}{T}\displaystyle\int_{-T/2}^{T/2} f(t)\mathrm{e}^{-\mathrm{i}n\Delta\omega t}\mathrm{d}t, \quad n \in \mathbf{Z}. \end{cases} \tag{3}$$

式(1)和式(2)叫做 Fourier 级数的三角形式,式(3)叫做其复指数形式,这两种形式之间的关系可由下面的关系式体现:

$$\begin{cases} C_0 = a_0, \\ |C_n| = |C_{-n}| = \dfrac{1}{2}\sqrt{a_n^2 + b_n^2} = a_{n/2}, \ n = 1, 2, \cdots, \\ \arg C_n = -\arg C_{-n} = \theta_n. \end{cases}$$

因此,C_n 作为一个复数,其模与辐角正好反映了函数(模拟信号)$f(t)$中频率为$n\Delta\omega$的谐波的振幅与相位. 同样称 C_n 为周期函数$f(t)$的离散频谱,$|C_n|$为离散振幅谱,$\arg C_n$为离散相位谱. Fourier 级数这两种表示形式的不同之处在于:其三角形式中仅用正频率的谐波表现周期函数 $f(t)$;而复指数形式中用正频率和负频率的谐波一起来表现 $f(t)$.

第二节　连续 Fourier 变换

前面讨论了周期函数可由离散频率为 $n\Delta\omega$ 的谐波来表现. 一个非周期函数 $f(t)$能否由谐波来表现呢?

一、Fourier 积分公式

将非周期函数 $f(t)$看作周期函数 $f_T(t)$当周期 $T\to+\infty$时转化的结果，于是，由本章第一节中式(3)有

$$\begin{aligned} f(t) &= \lim_{T\to+\infty} f_T(t) = \lim_{T\to+\infty}\sum_{n=-\infty}^{+\infty}\left[\frac{1}{T}\int_{-T/2}^{T/2} f_T(\tau)\mathrm{e}^{-\mathrm{i}n\Delta\omega\tau}\mathrm{d}\tau\right]\mathrm{e}^{\mathrm{i}n\Delta\omega t} \\ &= \frac{1}{2\pi}\lim_{\Delta\omega\to 0}\sum_{n=-\infty}^{\infty}\left[\left(\int_{-\pi/\Delta\omega}^{\pi/\Delta\omega} f_T(\tau)\mathrm{e}^{-\mathrm{i}n\Delta\omega\tau}\mathrm{d}\tau\right)\mathrm{e}^{\mathrm{i}n\Delta\omega t}\right]\Delta\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau\right]\mathrm{e}^{\mathrm{i}\omega t}d\omega. \end{aligned}$$

在以上推导中用到了本章第一节式(1)中的关系 $\Delta\omega = 2\pi/T$，还用到了关于和式极限的积分定义. 上式推导是有条件的，它允许非周期函数含有有限个间断点，间断点两侧的函数值是有界的，或即$f(t)$是有界变差的，对 $\forall t\in(-\infty,+\infty)$，有

$$f(t\pm 0) < +\infty.$$

上述推导还要求非周期函数 $f(t)$是绝对可积函数，即要求

$$\int_{-\infty}^{+\infty} |\ f(t)\ |\ \mathrm{d}t < +\infty.$$

这相当于对衰减性提出了较高的要求. 综上所述，记

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t, \tag{1}$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega, \tag{2}$$

对 $t\in(-\infty,+\infty)$，$f(t)$的谐波表示形式有如下的 Fourier 积分定理.

定理　对 $t\in(-\infty,+\infty)$，若 $f(t)$是有界变差且是绝对可积函数，则在 $f(t)$的连续点处，其函数值可用式(2)所示的 Fourier 积分表示，在$f(t)$的间断点处，$f(t) = [f(t-0)+f(t+0)]/2$.

二、Fourier 变换

在上述 Fourier 积分中，式(1)可看作积分变换，它将时间 t 作为自变量的时域函数 $f(t)$，$t\in(-\infty,+\infty)$，通过指定的积分运算，变换为频率 ω 作为自变量的频域函数 $F(\omega)$，$\omega\in(-\infty,+\infty)$；式(2)是式(1)所示变换的逆变换，它将频域函数$F(\omega)$通过类似的积分运算，变换为原来的时域函数$f(t)$；由式(1)和

式(2)所示的$f(t)$和$F(\omega)$是一一对应的变换对. 我们将其定义如下：

定义　式(1)叫做 Fourier 变换,记为 $FT[f(t)] = F(\omega)$, $F(\omega)$叫做$f(t)$的像函数;式(2)叫做 Fourier 逆变换,记为 $FT^{-1}[F(\omega)] = f(t)$, $f(t)$叫做$F(\omega)$的像原函数.

Fourier 变换和 Fourier 级数一样具有明确的物理意义. 由式(1)和式(2)可见,非周期函数$f(t)$也是由许多不同频率的谐波分量 $e^{i\omega t}$合成的. 不同的是,周期函数的谐波分量 $e^{in\Delta\omega t}$(本章第一节式(3))是离散分布的,非周期函数的谐波分量 $e^{i\omega t}$是连续分布的. 因此,$F(\omega)$反映了$f(t)$中各频率分量的分布密度,$F(\omega)$叫做频谱密度函数(简称频谱或连续频谱),$|F(\omega)|$叫做振幅谱,$\arg F(\omega)$叫做相位谱.

三、Fourier 变换与 Fourier 级数的联系

Fourier 级数是用离散频率 $n\Delta\omega$ 表现周期函数的,Fourier 变换是用连续频率 ω 表现非周期函数的,那么,在$n\Delta\omega$处级数和变换的表现结果有没有差别呢?

要作比较,首先要有可比性,对 Fourier 级数而言,它不仅能表现$[-T/2, T/2]$这个周期段的频谱分析,也能表现其他周期段的频谱分析;对 Fourier 变换而言,它不仅能表现周期函数的一个周期段的频谱分析,也能表现非周期函数的频谱分析. 因此,为了可比,选定 $f(t)$, $t \in [-T/2, T/2]$ 为函数段,选定$n\Delta\omega$作为比较频率.

一方面,可认为此函数段是某个周期函数的一个周期段,将其作周期延拓后用复指数形式的 Fourier 级数式(本章第一节式(3))表示,即知关于离散频率$n\Delta\omega$的谐波分量 $e^{-in\Delta\omega t}$ 的振幅为

$$C_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-in\Delta\omega t}\,dt.$$

另一方面,可将此函数段以外部分用零值代替,将其看作定义在$(-\infty, +\infty)$上的非周期函数,并使用式(1)和式(2),即任意频率 ω 所对应谐波分量 $e^{i\omega t}$的振幅为 $F(\omega)$,特别是对离散频率$n\Delta\omega$,其谐波分量的振幅为

$$\begin{aligned} F(\omega)\big|_{\omega=n\Delta\omega} &= \int_{-\infty}^{\infty} f(t)e^{-in\Delta\omega t}\,dt \\ &= \int_{-T/2}^{T/2} f(t)e^{-in\Delta\omega t}\,dt = TC_n. \end{aligned}$$

由此可见,就离散频率 $n\Delta\omega$ 可言,这两种分析的结果仅相差一个不变的常数,它们是等价的.

第三节　离散 Fourier 变换和谱函数的近似计算

一、离散 Fourier 变换及其快速算法

为了用计算机作频谱分析，需要把信号离散量（采样值）和频谱函数离散量作为 Fourier 变换对，建立对离散量的相应算法. 这就是离散 Fourier 变换.

设已知某函数段，它可以是周期函数的一个周期，也可以是非周期函数的主要部分段，表示为 $f(t)$，$t \in [-T/2, T/2]$. 用 Δt 间隔对 t 作等距分划，记 $f_j = f(j\Delta t)$，可获得函数段的样本值 $\{f_j\}_{j=-N}^{N}$，为了求得离散值 $\{F_k\}_{k=-N}^{N}$，按 Fourier 变换规则与样本值对应. 其中，$F_k \approx F(k\Delta\omega)$，$\Delta\omega$ 为连续频率变量 ω 的离散间隔. 于是，用数值积分的办法来模拟 Fourier 变换及其逆变换（本章第二节式(1)和式(2)），得

$$\begin{cases} f_j = \Delta\omega \sum\limits_{k=-N}^{N} F_k e^{i[j\Delta t][k\Delta\omega]}, \\ F_k = \Delta t \sum\limits_{j=-N}^{N} F_j e^{-i[j\Delta t][k\Delta\omega]}. \end{cases} \tag{1}$$

利用式(1)的互逆性要求，求得

$$\Delta\omega = 1/[(2N+1)\Delta t].$$

总之，相应于连续变量的 Fourier 变换对 $f(t) \sim F(\omega)$ 有离散的 Fourier 变换对 $\{f_j\}_{j=-N}^{N} \sim \{F_k\}_{k=-N}^{N}$.

从表面上看，按式(1)计算并不难，但仔细分析可知，要计算 N 个点的离散 Fourier 变换的工作量为 N^2 次复运算（一次复运算是指复数的一次乘法和加法）. 在实际应用中，要处理的信息量太大，并且要实时处理，需要运用快速算法，即快速 Fourier 变换算法（FFT 算法）. 它使计算工作量降为 $N\log_2 N$，提高工效倍数为 $N^2/N(\log_2 N) = N/\log_2 N$. 例如，当 $N = 10^3$ 时，提高工效约 100 倍，当 $N = 10^6$ 时提高工效约 50 000 倍. FFT 算法不仅计算速度快，而且计算精度高，计算过程稳定，它已被广泛地应用于谐波分析中.

二、取样效应

虽然 $\{f_j\}_{j=-N}^{N}$ 和 $\{F_k\}_{k=-N}^{N}$ 是离散 Fourier 变换对，但由于截取函数的有限长度和作等距离散的原因，用 $\{F_k\}$ 来表现 $F(\omega)$ 就不可避免地会产生误差.

一方面，$\{F_k\}$ 是有限长度的，它所表现的谐波的频率范围有限，有些高频量得不到反应.

事实上，在取样频率 $1/\Delta t$ 确定的情况下，某频率 s 都可表示为 $s = \omega_k + m/\Delta t$，其中，$m$ 是整数，ω_k 在$\{F_k\}$所能表现的频率范围内，于是在 Δt 间隔取样时$\{F_k\}$具体表现的谐波

$$e^{isj\Delta t} = e^{i[\omega_k + m/\Delta t]j\Delta t} = e^{i\omega_k j\Delta t}.$$

关于频率为 $m/\Delta t$ 的谐波分量是表现不出来的，无论怎样加工处理都无法分辨，它们被折合到比 $1/\Delta t$ 更低的频率 $|\omega_k| \leqslant 1/(2\Delta t)$ 的谐波中，这就是取样频率所造成的频率混叠效应. 换句话说，由$\{F_k\}$所表现的谐波频率范围为

$$|\omega_k| \leqslant \frac{1}{2\Delta t},$$

它与取样间隔 Δt 有关. 总之，由于取样间隔 Δt 不能无限小，于是由$\{F_k\}$所表现的频率辨认范围只能是有限大的，即 $|\omega_k| \leqslant 1/(2\Delta t)$. 这就表明，在应用中要根据需要分辨的频率范围来确定取样间隔Δt.

另一方面，由$\{F_k\}$所表现的频谱 $F(\omega)$被平滑. 在一定程度上破坏了 $F(\omega)$的精细结构.

事实上，在用离散 Fourier 变换作谐波分析时，通常是截取有限段 $-T/2 \leqslant t \leqslant T/2$ 来分析$f(t)$的，这就相当于待分析的函数是 $f(t)$，$t \in (-\infty, +\infty)$ 和 $rect(t) = \begin{cases} 1, 当 |t| \leqslant T/2 时; \\ 0, 当 |t| > T/2 时 \end{cases}$ 相乘所得的新函数

$$\tilde{f}(t) = f(t) \cdot rect(t) = \begin{cases} f(t), & 当 |t| \leqslant T/2 时; \\ 0, & 当 |t| > T/2 时. \end{cases}$$

于是，

$$FT[\tilde{f}(t)] = \tilde{F}(\omega) = FT[f(t)] * FT[rect(t)].$$

显然，$\tilde{F}$ 是 $F(\omega)$被平滑后的结果，它在一定程度上破坏了$F(\omega)$的精细结构.

三、谱的近似计算

要用离散 Fourier 变换近似计算谱函数，必须注意取样效应.

对时域函数(模拟信号)$f(t)$取有限宽度 T，应考虑到有效区段被截取；取采样间隔Δt，应考虑到时域中函数的精细变化被反映，这样，可获得时域的一串样本值 $\{f_j\}_{j=-N}^{N}$. 由取样效应分析可知，T 和 Δt 又决定了用$\{F_k\}$表现 $F(\omega)$的有效分辨精度和有效分辨范围，T 决定了精细分辨程度 $\Delta\omega = 1/T$，Δt 决定了有效的频率范围 $|\omega_k| \leqslant 1/(2\Delta t)$. 换言之，在离散 Fourier 变换中，用 $\{f_j\}_{j=-N}^{N}$ 来近似$f(t)$，时域分辨精度为Δt，时域分辨范围为 $-T/2 \leqslant t \leqslant T/2$，近似特点为 $f_j = f(t_j)$；相应的变换对 $\{F_k\}_{k=-N}^{N}$ 是近似$F(\omega)$的，有效的频域分辨精度为

$\Delta\omega = 1/T$，有效的频率分辨范围为 $|\omega_k| \leqslant 1/(2\Delta t)$，近似特点为 $F_k \approx F(\omega_k)$.

据此分析结果可确定近似计算谱函数的办法.

① 确定取样宽度 Δt，要求 $|\omega_k| \leqslant 1/(2\Delta t)$ 能包含实际需要分析的频率范围，同时兼顾对 $f(t)$ 变化的描述；

② 确定取样范围 T，或选定足够大的 N，要求按 $\Delta\omega = 1/T$ 分隔的离散频率 ω_k 处，$\{F_k\}$ 能反应 $F(\omega)$ 的波动变化；

③ 用 FFT 算法，由 $\{f_j\}_{j=-N}^{N}$ 计算出离散 Fourier 变换对 $\{F_k\}_{k=-N}^{N}$.

在实际应用中，常需要计算函数 $f(t)$ 的振幅谱 $|F(\omega)|$ 和能量密度函数 $|F(\omega)|^2$. 应该看到，利用 $\{F_k\}$ 计算这些所需要的离散近似值是不难的.

第四节　窗口 Fourier 变换

Fourier 分析在信号分析处理中的突出贡献，在于它将复杂的时域信号转换到频域中，用频谱特性去分析和表现时域信号的特性. 但 Fourier 分析不能分析局部时域信号的局部频谱特性，它没有时—频局部化功能. 窗口 Fourier 变换正是在时—频局部分析方面取得了本质性的进步. 关于时—频局部化概念、度量方法和时频窗的意义，对以后关于小波变换的学习是非常重要的.

一、短时的时—频分析需要

Fourier 分析能有效地分析平稳信号，能通过频谱函数方便地指明平稳信号的主要谐波成分. 然而，在实际应用中，我们常需要分析频域特性随时间变化的非平稳信号，如音乐信号、语音信号、探地信号等，需要了解某些局部时域信号所对应的主要频率特性，也需要了解某些频率的信息出现在哪些时间段上. 具体地说，对于音乐模拟信号，我们希望了解局部时域信号的频率特性，希望局部地改变音乐信号的听觉效果；对于基桩施工所采集的信号，我们希望知道什么局部时域信号表示基桩顶部达到什么性质的地层，什么局部时域信号表示基桩发生断裂现象；对于电网运行所采集的信号，我们希望实时地进行监测分析，快速判断故障信号发生的时间、地点和故障类型，以便能作出即时反应并保证电网正常运行. 上述情形都提出了关于短时段时域信号所对应的局部频域特性，即时—频局部化的要求.

Fourier 分析对时—频局部化要求是无能为力的. 只要仔细观察分析 Fourier 变换表达式

$$F(\omega) = \int_{\mathbf{R}} f(t)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t,\ \omega \in \mathbf{R}$$

即可明白其中的原因.一方面,Fourier 变换要求提供 $f(t)$, $t \in \mathbf{R}$ 的全部信息,即使截取短时段信号 $f(t)$, $t \in [-T/2, T/2]$, 但 $F(\omega)$ 提供的是关于 $\omega \in \mathbf{R}$ 的全部信息,而主要反映这个短时段时域信号的那些局部频域特性却无法知道.另一方面,时域信号 $f(t)$ 的局部改变会影响其 $F(\omega)$ 的全局改变, $F(\omega)$ 在某个特定的 ω^* 处的表现不可能通过局部时域信号得到,它需要提供 $f(t)$, $t \in R$ 的全部信息.特别值得注意的是, $FT[\delta(t)] = 1$, $FT[\delta(\omega)] = 1$, 这清楚地表明:时域中某点的局部变化会影响频域全局;同样,频域中的某点的局部变化也会影响到全部时域.因此,Fourier 分析方法没有时一频局部化功能.

二、窗口 Fourier 变换(WFT)的基本思想

为克服 Fourier 变换在时一频局部化方面的不足,D. Gabor 于 1946 年提出了窗口 Fourier 变换(简记为 WFT)方法.

WFT 的数学形式为

$$(Gf)(\omega, b) = \int_{\mathbf{R}} f(t) w(t-b) \mathrm{e}^{-\mathrm{i}\omega t} \mathrm{d}t, \tag{1}$$

其中,实函数 $w(t)$ 为时窗函数.由于式(1)右端形式是由 Gabor 提出的一种积分变换,此类变换称为 Gabor 变换,变换结果是关于 ω 和 b 的函数,故用记号 $(Gf)(\omega, b)$ 表示这种变换.

WFT(即式(1))实现时域和频域局部化的基本思想是重要的.在时域局部化方面,它通过引进的时窗函数 $w(t)$,使时域信号 $f(t)$ 在 $t=b$ 附近被局部化为 $f(t)w(t-b)$(图 6.1.1).在频域局部化表现方面,要观察 $(Gf)(\omega, b)$ 的表现,特别是对于某个确定的 $\omega = \eta$ 处关于 $(Gf)(\omega, b)$ 的表现.为此,利用卷积公式,有

$$\begin{aligned}
(Gf)(\omega, b) &= FT[f(t) \cdot w(t-b)] \\
&= FT[f(t)] * FT[w(t-b)] \\
&= \hat{f}(\omega) * (\hat{w}(\omega) \mathrm{e}^{-\mathrm{i}\omega b}) \\
&= \int_{\mathbf{R}} \hat{f}(\xi)\, \bar{\hat{w}}(\xi - \omega) \mathrm{e}^{\mathrm{i}(\xi-\omega)b} \mathrm{d}\xi \\
&= 2\pi \mathrm{e}^{-\mathrm{i}\omega b} FT^{-1}[\hat{f}(\xi)\, \bar{\hat{w}}(\xi - \omega)].
\end{aligned} \tag{2}$$

其中, $\hat{f}(\omega) = FT[f(t)]$, $\hat{w}(\omega) = L[w(t)]$ 是通用记号.由以上两式可清楚地看出,若 $\hat{w}(\omega)$ 在 $\omega = \eta$ 附近是有局部化作用的,频域信息 $\hat{f}(\omega)$ 就能在 $\omega = \eta$ 附近被局部化.总之,粗略地看,WFT 是在 $t = b$ 附近的时间窗中观察时域信号 $f(t)$,在 $\omega = \eta$ 附近的频率窗中观察频域信号 $\hat{f}(\omega)$.显然,要使 WFT 能用于时一频局部化分析,必须对 $w(t)$ 和 $\hat{w}(\omega)$ 提出合适的要求.

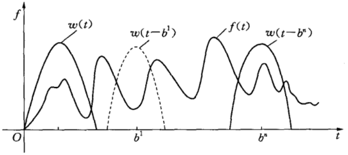


图 6.1.1 时窗函数 $w(t)$的时域局部化表现

三、时窗、频窗和时一频窗

1. 时窗及其度量

时窗函数起着时域局部化的作用，$f(t)w(t-b)$ 就表明了 $t=b$ 处的局部时域信号. 时窗函数 $w(t)$ 的开窗效果是用时窗中心 $t^*[w(t)]$ 和时窗半径 $\Delta_t[w(t)]$来表示的，若无须指明窗函数，可直接使用记号 t^* 和 Δ_t. 时窗中心可依照力学中关于重心的描述来定义，即

$$t^* = \int_{\mathbf{R}} t \mid w(t) \mid^2 \mathrm{d}t \Big/ \int_{\mathbf{R}} \mid w(t) \mid^2 \mathrm{d}t.$$

时窗半径可依照力学中求矩量的办法来定义，即

$$\Delta_t = \left\{ \int_{\mathbf{R}} (t-t^*)^2 \mid w(t) \mid^2 \mathrm{d}t \Big/ \int_{\mathbf{R}} \mid w(t) \mid^2 \mathrm{d}t \right\}^{1/2}$$

特别地，令 $\int_{\mathbf{R}} \mid w(t) \mid^2 \mathrm{d}t = 1$，这可以通过调整 $w(t)$的形状简单地做到. 此时，时窗函数 $w(t-b)$ 当 $b=0$ 时的时窗中心和时窗半径定义描述为

$$\begin{cases} t^* = \int_{\mathbf{R}} t \mid w(t) \mid^2 \mathrm{d}t, \\ \Delta_t = \left\{ \int_{\mathbf{R}} (t-t^*)^2 \mid w(t) \mid^2 \mathrm{d}t \right\}^{1/2} \end{cases} \tag{3}$$

在此定义下，时窗函数 $w(t-b)$（其中 $b=0$）的限时作用表现为以 t^* 为中心的 $[t^*-\Delta_t, t^*+\Delta_t]$ 范围，时窗宽度为$2\Delta_t$.

同样，在此定义下，对于时窗函数 $w(t-b)$（其中 $b\neq 0$）的情形，利用时窗中心和时窗半径的直观几何含义，不难知道

$$\begin{cases} t^*[w(t-b)] = t^*[w(t)] + b, \\ \Delta_t[w(t-b)] = \Delta_t[w(t)]. \end{cases} \tag{4}$$

事实上,数学推导也是成立的,利用式(3),有

$$
\begin{aligned}
t^*[w(t-b)] &= \int_{\mathbf{R}} t \mid w(t-b) \mid^2 \mathrm{d}t \\
&= \int_{\mathbf{R}} (x+b) \mid w(x) \mid^2 \mathrm{d}x \\
&= t^*[w(t)] + b, \\
\mid \Delta_t[w(t-b)] \mid^2 &= \int_{\mathbf{R}} [t-(t^*+b)]^2 \mid w(t-b) \mid^2 \mathrm{d}t \\
&= \int_{\mathbf{R}} [x+b-(t^*+b)]^2 \mid w(x) \mid^2 \mathrm{d}x \\
&= \int_{\mathbf{R}} \mid t-t^* \mid^2 \mid w(t) \mid^2 \mathrm{d}t \\
&= \Delta_t[w(t)]^2.
\end{aligned}
$$

2. 频窗及其度量

在 WFT 中,$w(t)$起着时域局部化的作用,$\hat{w}(\omega)$起着频域局部化的作用,$w(t)$是时窗函数,$\hat{w}(\omega)$是频窗函数.故频窗函数的开窗效果度量类似于时窗效果度量(见式(3)).

频窗中心 ω^* 和频窗半径 Δ_ω 定义为

$$
\begin{cases}
\omega^* = \int_{\mathbf{R}} \omega \mid \hat{w}(\omega) \mid^2 \mathrm{d}\omega \Big/ \int_{\mathbf{R}} \mid \hat{w}(\omega) \mid^2 \mathrm{d}\omega, \\
\Delta_\omega = \left\{ \int_{\mathbf{R}} (\omega - \omega^*)^2 \mid \hat{w}(\omega) \mid^2 \mathrm{d}\omega \Big/ \int_{\mathbf{R}} \mid \hat{w}(\omega) \mid^2 \mathrm{d}\omega \right\}^{1/2}.
\end{cases} \tag{5}
$$

同时窗一样,频窗平移 η 后,有

$$
\begin{cases}
\omega^*[\hat{w}(\omega-\eta)] = \omega^* + \eta, \\
\Delta_\omega[\hat{w}(\omega-\eta)] = \Delta_\omega.
\end{cases} \tag{6}
$$

3. 作为窗函数的条件

如果 $w(t)$作为时窗函数,不一定要求它是紧支的(即非零函数值是有限范围的),它可以是定义在 **R** 上的实函数,关键是要满足式(3)中的 t^* 和 Δ_t 是有限的,即要求

$$
\int_{\mathbf{R}} t^2 \mid w(t) \mid^2 \mathrm{d}t < +\infty.
$$

据 **R** 上定积分存在判别法则,要求当 $\mid t \mid \to +\infty$ 时,$\mid t \mid^2 \mid w(t) \mid^2 < 1/\mid t \mid$,即

$$
\mid w(t) \mid < 1/\mid t \mid^{3/2}, \quad \mid t \mid \to +\infty, \tag{7}
$$

这就是说，当$|t|\to+\infty$时，$|w(t)|$的衰减速度应比$1/|t|^{3/2}$快，这样的函数可作时窗函数.

如果要求$\hat{w}(\omega)$作频窗函数，仿照前面的讨论，$\hat{w}(\omega)$应使得ω^*和Δ_ω是有限值，也要求

$$|\hat{w}(\omega)|<1/|\omega|^{3/2},\ |\omega|\to+\infty. \tag{8}$$

总之，一个函数$w(t)$，要使其作为时窗函数，其谱函数作为频窗函数，则$w(t)$和$\hat{w}(\omega)$应同时具有较强的衰减性(见式(7)和式(8)).

4. 时—频窗

若$w(t)$和$\hat{w}(\omega)$能分别起着时窗和频窗的作用，则可简称$w(t)$为窗函数. 它的时域局部化作用被限制在时窗区间$[(t^*+b)-\Delta_t,\ (t^*+b)+\Delta_t]$范围内；它的频域局部化作用被局限在频窗区间$[(\omega^*+\eta)-\Delta_\omega,\ (\omega^*+\eta)+\Delta_\omega]$范围内. 于是，可构建时—频坐标系(图 6.1.2)，用时窗区间和频窗区间形成一个矩形时—频窗. 时—频窗是窗函数的时—频局部化功能的几何直观的描述.

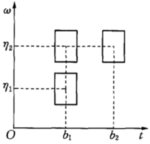


图 6.1.2 WFT 时窗示意图

另外，由式(4)和式(6)可知，在 WFT 中，在时窗中心和频窗中心有移动的情形下，其时窗半径Δ_t和Δ_ω是不会改变的. 换句话说，在 WFT 的时—频坐标系中，任意的窗中心$(b,\ \eta)$所在的时频窗面积恒为$4\Delta_t\Delta_\omega$不变.

5. Gauss 窗函数

前面已讨论了窗函数所必须满足的条件式(7)和式(8)，简言之，窗函数在时域和频域中都具有快速衰减性. 很多函数都可以作为窗函数，例如矩形函数、山形函数、m次样条函数等.

这里重点介绍 Gauss 窗函数. 取用 Gauss 函数的变形形式

$$g_a(t)=\frac{1}{2\sqrt{\pi a}}\mathrm{e}^{-t^2/(4a)},\quad a>0$$

作为时窗函数，$g_a(t)$是偶对称且是快速衰减的. 可以选择实常数$a>0$来实现函数的伸缩，调整时窗的宽窄；也可以选择实常数b，使$g_a(t-b)$实现函数平移，调整时窗的位置. 因为

$$\hat{g}_a(\omega)=\frac{1}{4a}\mathrm{e}^{-a\omega^2},\ a>0$$

仍是 Gauss 函数，也是偶对称和快速衰减的，所以$\hat{g}_a(\omega)$又可作频窗函数.

可方便地看出，其时窗中心和频窗中心为

$$t^*[g_a(t)] = 0, \quad \omega^*[\hat{g}_a(\omega)] = 0.$$

为了计算时窗和频窗半径，要用到公式

$$\begin{cases} \int_{\mathbf{R}} e^{-\xi t^2} dt = \pi^{1/2}\xi^{1/2}, \\ \int_{\mathbf{R}} t^2 e^{-\xi t^2} dt = \frac{1}{2}\pi^{1/2}\xi^{-3/2}. \end{cases}$$

只要令 $\xi = 1/(4a)$，就有

$$\int_{\mathbf{R}} [g_a(t)]^2 dt = (8\pi a)^{-1/4},$$

$$\int_{\mathbf{R}} t^2 [g_a(t)]^2 dt = \left[\frac{1}{4\pi a} \cdot \frac{\sqrt{\pi}}{2}(2a)^{3/2}\right]^{1/2}.$$

再按式(5)算出

$$\Delta_t[g_a(t)] = \sqrt{a}.$$

同理，令 $\xi = a$, $t = \omega$，即可算出

$$\Delta_\omega[\hat{g}_a(\omega)] = 1/(2\sqrt{a}).$$

于是，当采用 $g_a(t-b)$ 作窗函数时，

时—频窗中心为 $O = (b, \eta)$，

时—频窗位置为 $[b - \Delta_t, b + \Delta_t] \times [\eta - \Delta_\omega, \eta + \Delta_\omega]$，

时—频窗面积为 $A = (2\Delta_t)(2\Delta_\omega) = 2.$

$g_a(t-b)$ 作为窗函数时，相应的 WFT 在 $\omega - t$ 相平面中的时—频窗示意图如图 6.1.2 所示.

在 WFT 中，用 $g_a(t)$作窗函数和其他的$w(t)$作窗函数一样，无论时—频窗中心(b, η)处于何处，时—频窗总是形状相同且面积相等的.

四、WFT 反演公式

任何一种积分变换，只有当它具有反演公式时，才会有意义.

相应于式(1)WFT 的反演公式为

$$f(t) = \frac{1}{2\pi}\int_{\mathbf{R}}\int_{\mathbf{R}} [e^{i\omega t}(Gf)(\omega, b)] w(t-b) d\omega db, \tag{9}$$

其中，函数 $f(t)$经 WFT 得到$(Gf)(\omega, b)$，它是关于连续变量 $\omega \in R$ 和$b \in R$ 的函数，故式(9)中会出现对 $\omega \in R$ 和 $b \in R$ 的积分.

式(9)的证明是不难的. 事实上,先把式(1)中的 $f(t)w(t-b)$ 看成整体形式,再对 $(Gf)(\omega, b)$ 作 Fourier 逆变换,有

$$f(t)w(t-b) = \frac{1}{2\pi}\int_{\mathbf{R}} e^{i\omega t}(Gf)(\omega, b)d\omega,$$

上式两边同时乘以 $w(t-b)$,再对变量 b 积分. 注意,对实函数 $w(t-b)$ 总可假定

$$\int_{\mathbf{R}}[w(t)]^2 dt = \int_{\mathbf{R}}[w(t-b)]^2 db = 1, \tag{10}$$

这样就可以看出式(9)的正确性.

类似于 Fourier 变换的乘积定理和 Parseval 等式,WFT 也有相应的公式

$$\int_{\mathbf{R}}\int_{\mathbf{R}}(Gf)(\omega, b)\overline{(Gg)(\omega, b)}d\omega db = 2\pi\int_{\mathbf{R}} f(t)g(t)dt, \tag{11}$$

$$\int_{\mathbf{R}}\int_{\mathbf{R}} |(Gf)(\omega, b)|^2 d\omega db = 2\pi\int_{\mathbf{R}}[f(t)]^2 dt, \tag{12}$$

其中,$f(t)$、$g(t)$和 $w(t)$都是实函数.

下面证明式(11).

先把 WFT 看做某函数的 Fourier 变换,即

$$(Gf)(\omega, b) = FT[f(t)w(t)],$$

于是利用 Parseval 等式,有

$$\begin{aligned}
&\int_{\mathbf{R}}(Gf)(\omega, b)\overline{(Gg)(\omega, b)}d\omega \\
&= \int_{\mathbf{R}} FT[f(t)w(t-b)]\overline{FT[g(t)w(t-b)]}d\omega \\
&= 2\pi\int_{\mathbf{R}}[f(t)w(t-b)]\overline{[g(t)w(t-b)]}d\omega \\
&= 2\pi\int_{\mathbf{R}} f(t)g(t)[w(t-b)]^2 dt,
\end{aligned}$$

再将上式对变量 b 积分,并利用式(10),立即可证得式(11).

在式(11)中令 $g(t) = f(t)$,即有式(12).

五、WFT 的某些局限性

就时—频局部化而言,WFT 在 Fourier 分析的基础上取得了本质的进步. 用 WFT 分析信号可在时—频窗这个局部范围内观察,时—频窗面积反映了时—频局部化的精细程度. 是否可以选择某个窗函数,能使时—频窗面积充分小

呢？Heissenberg 测不准原理表明，任何窗函数所相应的时—频窗面积都有 $A \geqslant 2$. 这就是说，就时—频窗面积而言，Gauss 窗函数 $g_a(t)$ 已经是最好的结果了.

时—频局部化的精细程度还反映在时—频窗形状上.低频信号的特点是，大的时间范围内幅值变化慢，其频率范围窄，于是分析低频信号的时—频窗特点应是时窗宽且频窗窄；高频信号的特点是较短的时间范围内幅值变化快，其频率范围宽，于是分析高频信号的时—频窗特点应是时窗窄且频窗宽. WFT 分析信号的时—频窗形状有如下特点：以 Gauss 窗函数 $g_a(t-b)$ 为例，选定 a，其时—频窗形状确定；若 $f(t)w(t-b_1)$ 和 $f(t)w(t-b_2)$ 相同时，被局部化的信号的频带和频率中心不变，则不变形状的时频窗从 b_1 处移至 b_2 处；若时窗 $w(t-b_1)$ 不变，信号发生变化，则 $f(t)w(t-b_1)$ 和 $\hat{f}(t)w(t-b_1)$ 的中心频率将由 η_1 处移至 η_2 处，此时的时—频窗形状仍然不变(图 6.1.2)；为了分析局部低频信号，可选择较大的参数 a，使 Δ_t 增大，Δ_ω 变窄；为了分析局部高频信号，可再选择较小的参数 a，使 Δ_t 变窄，Δ_ω 变宽.总之，虽然 WFT 可分析低频和高频信号，但其表现显得呆滞；若用确定 a 的时—频窗来分析同一信号的不同时段，则有时适合，有时不适合；若通过调整参数 a 来分析同一信号的不同时段，则显得很麻烦.

时—频局部化的精细程度还表现于分析高、低频信号的自适应时—频窗方面.在给定的实际模拟信号中，常常同时含有表示某种特定情况的低频信号段和表示另外特定情况的高频信号段，此时需要一种自适应的时—频局部化方法.也就是说，选择某个窗函数，希望其时—频窗形状是自适应变化的，对低频信号，其窗口形状自动变得扁平，对高频信号，其窗口形状自动变得瘦长.对于这方面的要求，WFT 是无能为力的.

在时—频局部分析中往往还要求将频域分成若干个细小频带，进一步分析信号在各个频带中的表现，往往还要实时地、精细程度更高地、自适应地分析信号.对于这些要求，WFT 也是无能为力的.下面介绍的小波分析方法对满足这些要求做出了突出的贡献.

第二章　小 波 变 换

WFT 在时频分析中，不能根据高、低频信号的特点，自适应地调整时—频窗，它在时—频局部化的精细方面和灵活方面表现欠佳. 本小节介绍的小波变换可以克服这一缺点. 针对自适应时—频窗的要求，本小节从积分变换出发，先分析这种窗函数应具备的抽象形式，再分析这类窗函数应满足的条件.

第一节　自适应窗函数的设计

正如前述所讨论的那样，在实际的时—频局部分析中，需要一种自适应的时—频窗. 在分析低频信号时，它的时窗宽而频窗窄；在分析高频信号时，它的时窗窄而频窗宽. 什么形式的窗函数才能适应这些要求呢？

为此，先分析 WFT 不能实现自适应时—频窗的原因. 在 WFT 中，通常是以时域开窗性能为主来考虑问题的，也就是说，

$$(Gf)(\omega,\ b) = \int_{\mathbf{R}} [f(t)w(t-b)]\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t.$$

先将 $f(t)$时域局部化为 $f(t)w(t-b)$，再对开窗后局部时域信号作 Fourier 变换. 在这种观点下，$w(t-b)$ 的设计难以自动适应低、高频信号在时域和频域中的局部表现，从而出现了时—频形状不变的特点. 也可换一个角度观察 WFT，即采用形式

$$(Gf)(\omega,\ b) = \int_{\mathbf{R}} f(t)\ \overline{\tilde{w}(\omega,\ t-b)}\,\mathrm{d}t,$$

其中，$\tilde{w}(\omega,\ t-b) = w(t-b)\mathrm{e}^{\mathrm{i}\omega t}$. 这是一种新的思考方式，在积分变换的意义下，既把 $\tilde{w}$ 看做变换函数，又把 $\tilde{\omega}$ 看做对$f(t)$在时域和频域能起限制作用的窗函数. 基于这种考虑，不妨假设窗函数都具有抽象形式 $\tilde{w} = \Psi_{ab}(t)$，其中

$$\Psi_{ab}(t) = a^{1/2}\Psi(at-b),\ a \in \mathbf{R}_{+},$$

它是由 $\Psi(t)$经平移和放缩的结果(图 6.2.1).

这种形式的窗函数能同时表现时间和频率方面的特征. 若 $a < 1$，则 $\Psi(t)$ 被拉宽且振幅被压低，$\Psi_{ab}(t)$含有表现低频分量的特征；若 $a > 1$，则$\Psi(t)$被压窄且振幅拉高，$\Psi_{ab}(t)$含有表现高频分量的特征. 因为 a 在$\Psi_{ab}(t)$中作为表现频率的参数，所以它不仅能适应关于不同频率时域信号的时窗函数的要求，而且在

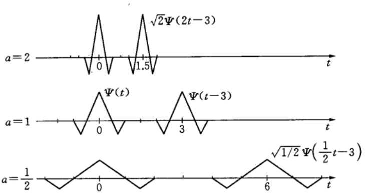


图 6.2.1 $\Psi_{ab}(t)$是 $\Psi(t)$经平移和伸缩的结果(其中 $b=3$)

$\hat{\Psi}_{ab}(\omega)$中也含有参数 a，$\hat{\Psi}_{ab}(\omega)$也能适应关于不同频率的频窗函数的要求. 对这些特点，下面会有定量描述.

第二节 小波、小波变换的定义与条件

根据前面的分析，对模拟信号 $f(t)$的积分变换

$$W_f(a,\ b)=\int_{\mathbf{R}} f(t)\overline{\Psi}_{ab}(t)\,\mathrm{d}t \tag{1}$$

叫做小波变换，其中，

$$\Psi_{ab}(t)=|\ a\ |^{1/2}\Psi(at-b) \tag{2}$$

是由 $\Psi(t)$经平移和放缩的结果. 当满足一定条件时 $\Psi(t)$叫做允许小波函数. 由式(2)有

$$\begin{aligned}\|\ \Psi_{ab}(t)\ \|_0^2&=\int_{\mathbf{R}}|\ \Psi_{ab}(t)\ |^2\mathrm{d}t\\&=\int_{\mathbf{R}}|\ \Psi(t)\ |^2\mathrm{d}t=\ \|\ \Psi(t)\ \|_0^2,\end{aligned} \tag{3}$$

$$\begin{aligned}FT[\Psi_{ab}(t)]&=\hat{\Psi}_{ab}(\omega)\\&=\int_{\mathbf{R}} a^{1/2}\Psi(at-b)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t\\&=a^{-1/2}\int_{\mathbf{R}}\Psi(at-b)\mathrm{e}^{-\mathrm{i}\frac{\omega}{a}(at-b)}\,\mathrm{e}^{-\mathrm{i}\omega\frac{b}{a}}\,\mathrm{d}(at-b)\\&=a^{-1/2}\mathrm{e}^{-\mathrm{i}\omega\frac{b}{a}}\hat{\Psi}\left(\frac{\omega}{a}\right).\end{aligned} \tag{4}$$

其中 $\| \Psi(t) \|_{0}^{2}$ 表明了 $\Psi(t)$ 的一种模量，式(3)说明，$\Psi(t)$ 经如式(2)那样作平移和放缩变化之后，其模量是不变的. 式(4)则体现出 $\hat{\Psi}_{ab}(\omega)$ 与参数 a 关联，且会随着 a 的变化而改变.

小波变换作为一种积分变换，只有当它能作回复变换时，它才是有意义的. 为了推导小波变换的回复公式，参照 Fourier 积分变换推导回复公式所需要的乘积定理式：

$$\frac{1}{2\pi}\int_{\mathbf{R}} FT[f(t)]\,\overline{FT}[g(t)]\mathrm{d}\omega = \int_{\mathbf{R}} f(t)\,\overline{g(t)}\,\mathrm{d}t.$$

上式表明了变换前两个函数的能量积分和它们在变换后的能量积分之间的等价关系，而且，当 $g(t) = \delta(t-x)$ 时，由上式可推导出 Fourier 逆变换公式. 受此启发，当积分变换是小波变换时，可以得到类似于乘积定理式的表现形式.

先观察

$$\begin{aligned} W_f(a,\ b) &= \int_{\mathbf{R}} f(t)\,\overline{\Psi}_{ab}(t)\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{\mathbf{R}} FT[f(t)]\,\overline{FT}[\overline{\Psi}_{ab}(t)]\mathrm{d}\omega \\ &= \frac{a^{-1/2}}{2\pi}\int_{\mathbf{R}} \hat{f}(\omega)\,\overline{\hat{\Psi}}\left(\frac{\omega}{a}\right)\mathrm{e}^{\mathrm{i}\frac{b}{a}\omega}\,\mathrm{d}\omega \\ &= \frac{a^{1/2}}{2\pi}\int_{\mathbf{R}} \hat{f}(a\omega)\,\overline{\hat{\Psi}}(\omega)\,\mathrm{e}^{\mathrm{i}b\omega}\,\mathrm{d}\omega \\ &= \frac{a^{1/2}}{2\pi}\,\overline{FT}[\overline{\hat{f}}(a\omega)\hat{\Psi}(\omega)], \end{aligned}$$

结果是变量 a 的函数. 上式推导中用到了积分变换的共轭关系，例如

$$\int_{\mathbf{R}} \xi(\omega)\,\mathrm{e}^{\mathrm{i}b\omega}\,\mathrm{d}\omega = \int_{\mathbf{R}} \overline{[\bar{\xi}(\omega)\,\mathrm{e}^{-\mathrm{i}b\omega}]}\mathrm{d}\omega = \bar{\xi}[\bar{\xi}(\omega)].$$

比照推导 $W_f(a,\ b)$ 的同样办法，有

$$\begin{aligned} W_g(a,\ b) &= \int_{\mathbf{R}} g(t)\,\overline{\Psi}_{ab}(t)\,\mathrm{d}t \\ &= \frac{a^{1/2}}{2\pi}\,\overline{FT}[\bar{\hat{g}}(a\omega)\hat{\Psi}(\omega)]. \end{aligned}$$

下面再观察

$$\begin{aligned} &\int_{\mathbf{R}} W_f(a,\ b)\,\overline{W_g}(a,\ b)\,\mathrm{d}b \\ =& \int_{\mathbf{R}} \left\{\frac{a^{1/2}}{2\pi}\,\overline{FT}[\overline{\hat{f}}(a\omega)\overline{\Psi}(\omega)]\right\}\left\{\frac{a^{1/2}}{2\pi} FT[\bar{\hat{g}}(a\omega)\hat{\Psi}(\omega)]\right\}\mathrm{d}b \end{aligned}$$

$$= \frac{a}{2\pi} \cdot \frac{1}{2\pi}\int_{\mathbf{R}} \overline{FT}[\overline{\hat{f}}(a\omega)\hat{\Psi}(\omega)]FT[\bar{\hat{g}}(a\omega)\hat{\Psi}(\omega)]\mathrm{d}b$$

$$= \frac{a}{2\pi}\int_{\mathbf{R}}[\hat{f}(a\omega)\overline{\hat{\Psi}}(\omega)][\bar{\hat{g}}(a\omega)\hat{\Psi}(\omega)]\mathrm{d}\omega.$$

利用上式结果，又有

$$\int_{\mathbf{R}}\left\{\int_{\mathbf{R}} W_f(a, b)\overline{W}_g(a, b)\mathrm{d}b\right\}\frac{1}{a}\mathrm{d}a$$

$$= \int_{\mathbf{R}}\left\{\frac{a}{2\pi}\int_{\mathbf{R}}\hat{f}(a\omega)\bar{\hat{g}}(a\omega)\hat{\Psi}(\omega)\overline{\hat{\Psi}}(\omega)\mathrm{d}\omega\right\}\frac{1}{a}\mathrm{d}a$$

$$= \left\{\int_{\mathbf{R}}\frac{|\hat{\Psi}(\omega)|^2}{\omega}\mathrm{d}\omega\right\}\frac{1}{2\pi}\int_{\mathbf{R}}\hat{f}(a\omega)\bar{\hat{g}}(a\omega)\mathrm{d}(a\omega)$$

$$= C_\Psi\frac{1}{2\pi}\int_{\mathbf{R}}\hat{f}(\xi)\bar{\hat{g}}(\xi)\mathrm{d}\xi$$

$$= C_\Psi\int_{\mathbf{R}} f(t)\bar{g}(t)\mathrm{d}t.$$

故小波变换有类似于 Fourier 变换中乘积定理式的结果：

$$\begin{cases}\int_{\mathbf{R}}\left\{\int_{\mathbf{R}} W_f(a, b)\overline{W}_g(a, b)\mathrm{d}b\right\}\frac{1}{a}\mathrm{d}a = C_\Psi\int_{\mathbf{R}} f(t)\bar{g}(t)\mathrm{d}t, \\ C_\Psi = \int_{\mathbf{R}}\frac{|\hat{\Psi}(\omega)|^2}{\omega}\mathrm{d}\omega < +\infty.\end{cases} \tag{5}$$

在式(5)中，取 $g(t) = f(t)$ 就得到类似于 Parseval 等式的能量关系式；在式(5)中取 $g(t) = \delta(t-x)$，就有

$$W_g(a, b) = \int_{\mathbf{R}}\delta(t-x)\overline{\Psi}_{ab}(t)\mathrm{d}t = \overline{\Psi}_{ab}(x),$$

其中用到公式

$$\int_{\mathbf{R}} f(t)\delta(t-x)\mathrm{d}t = f(x),$$

故由式(5)可推得回复公式

$$\int_{\mathbf{R}}\left\{\int_{\mathbf{R}} W_f(a, b)\Psi_{ab}(t)\mathrm{d}b\right\}\frac{1}{a}\mathrm{d}a = C_\Psi f(t). \tag{6}$$

不难理解，式(5)和式(6)成立的条件是小波变换被允许的依据. 由此可进一步分析允许小波函数$\Psi(t)$所必须具备的条件.

下面分析允许小波函数 $\Psi(t)$应满足的条件. 要使式(1)的积分变换有意义，一方面，由式(3)要求 $\int_{\mathbf{R}}|\Psi(t)|^2\mathrm{d}t < +\infty$，这就要求$\Psi(t)$具有快速衰减性；另

一方面，由$\hat{\Psi}(\omega)$的连续性和$|C_\Psi|<+\infty$可推知$\hat{\Psi}(0)=0$，由 Fourier 变换的表达式可推知

$$\int_{\mathbf{R}}\Psi(t)\mathrm{d}t=0.$$

这表明$\Psi(t)$具有波动性. 总之，$\Psi(t)$是像波一样的快速衰减函数，形如小的波，这也就是$\Psi(t)$称为小波的原因.

由于在$\int_{\mathbf{R}}\Psi(t)\mathrm{d}t=0$的条件下，$|C_\Psi|<+\infty$和衰减性表现

$$|\Psi(t)|<c(1+|t|)^{-1-\varepsilon},\ \varepsilon>0$$

是等价的，故关于允许小波的定义可采用下面两种等价的形式：

① 满足条件$|C_\Psi|<+\infty$的函数$\Psi(t)\in L^2(R)$叫做允许小波；

② 满足波动性$\int_{\mathbf{R}}\Psi(t)\mathrm{d}t=0$和衰减性$|\Psi(t)|\leqslant c(1+|t|)^{-1-\varepsilon}(\varepsilon>0)$的函数叫做允许小波.

应该注意，允许小波只是允许可作正向和逆向的小波变换的函数，若要让$\Psi(t)$和$\Psi_{ab}(t)$能作窗函数，还必须满足关于窗中心和窗半径的度量要求. 所以，可作窗函数的小波函数除保持振荡性以外，其衰减速度要比允许小波的快；作为窗函数的小波函数$\Psi_{ab}(t)$，是由同一函数$\Psi(t)$经平移和放缩得到的，$\Psi_{ab}(t)$在时域和频域都具有较强的局部化功能.

第三节　小波变换的自适应时—频窗

在小波变换的定义中，小波函数$\Psi_{ab}(t)$是窗函数，它的时—频窗表现了小波变换的时—频局部化能力. 依照前面的描述和讨论，本小节将主要介绍$\Psi_{ab}(t)$的开窗效果.

记t^*为时窗中心，Δ_t为时窗半径，ω^*为频窗中心，Δ_ω为频窗半径，则关于窗函数$\Psi_{ab}(t)$，有

$$\begin{cases} t^*=\dfrac{1}{\|\Psi_{ab}(t)\|_0^2}\displaystyle\int_{\mathbf{R}}t\,|\Psi_{ab}(t)|^2\mathrm{d}t, \\ \Delta t=\dfrac{1}{\|\Psi_{ab}(t)\|_0}\left\{\displaystyle\int_{\mathbf{R}}(t-t^*)^2\,|\Psi_{ab}(t)|^2\mathrm{d}t\right\}^{1/2}, \\ \omega^*=\dfrac{1}{\|\hat{\Psi}_{ab}(\omega)\|_0^2}\displaystyle\int_{\mathbf{R}}\omega\,|\hat{\Psi}_{ab}(\omega)|^2\mathrm{d}\omega, \\ \Delta_\omega=\dfrac{1}{\|\hat{\Psi}_{ab}(\omega)\|_0}\left\{\displaystyle\int_{\mathbf{R}}(\omega-\omega^*)^2\,|\hat{\Psi}_{ab}(\omega)|^2\mathrm{d}\omega\right\}^{1/2}. \end{cases} \tag{1}$$

关于窗函数 $\Psi(t)$ 的度量，可在式(1)中令 $a=1$、$b=0$ 得到，因为这些是基本量，所以特别标记为 t_{Ψ}^*、Δ_{Ψ}、$\omega_{\hat{\Psi}}^*$ 和 $\Delta_{\hat{\Psi}}$.

由于 $\Psi_{ab}(t)$ 是由 $\Psi(t)$ 经平移和放缩后的结果，所以式(1)中关于 $\Psi_{ab}(t)$ 的窗的度量可用 $\Psi(t)$ 的相应量表示. 因为

$$
\begin{aligned}
\| \Psi_{ab}(t) \|_0^2 &= \int_{\mathbf{R}} | \Psi_{ab}(t) |^2 \mathrm{d}t \\
&= \int_{\mathbf{R}} | a^{1/2} \Psi(at-b) |^2 \mathrm{d}t \\
&= \int_{\mathbf{R}} | \Psi(at-b) |^2 \mathrm{d}(at-b) \\
&= \| \Psi(t) \|_0^2 = \| \Psi(x) \|_0^2,
\end{aligned} \tag{2}
$$

所以，利用变量代换有

$$
\begin{aligned}
t^* &= \frac{1}{\| \Psi_{ab}(t) \|_0^2} \int_{\mathbf{R}} t \, | \Psi_{ab}(t) |^2 \mathrm{d}t \\
&= \frac{1}{\| \Psi(x) \|_0^2} \int_{\mathbf{R}} \left(\frac{x+b}{a} \right) | \Psi(x) |^2 \mathrm{d}x \\
&= \frac{1}{a} t_{\Psi}^* + \frac{b}{a} \\
&= \frac{1}{a} (t_{\Psi}^* + b),
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
\Delta_t &= \frac{1}{\| \Psi_{ab}(t) \|_0} \left\{ \int_{\mathbf{R}} (t-t^*)^2 \, | \Psi_{ab}(t) |^2 \mathrm{d}t \right\}^{1/2} \\
&= \frac{1}{\| \Psi_{ab}(t) \|_0} \left\{ \int_{\mathbf{R}} \left(t - \frac{1}{a} t_{\Psi}^* - \frac{b}{a} \right)^2 | \Psi(at-b) |^2 \mathrm{d}(at-b) \right\}^{1/2} \\
&= \frac{1}{a} \frac{2}{\| \Psi(x) \|_0} \left\{ \int_{\mathbf{R}} (x - t_{\Psi}^*)^2 \, | \Psi(x) |^2 \mathrm{d}x \right\}^{1/2} \\
&= \frac{1}{a} \Delta_{\Psi}.
\end{aligned} \tag{4}
$$

式(3)表明，若 $\Psi(t)$ 不被压缩，$\Psi(t)$ 仅向右移动 b 而变为 $\Psi(t-b)$，则 $\Psi(t)$ 的时窗中心 t_{Ψ}^* 也被向右移动 b，变为 t_{Ψ}^*+b；若 $\Psi(t)$ 变为 $\Psi_{ab}(t)$，则时窗中心就由 t_{Ψ}^* 变为 $\frac{1}{a}(t_{\Psi}^*+b)$. 式(4)表明，对 $\Psi(t)$ 的平移并不改变其时窗宽度；当 $\Psi(t)$ 被压缩 $1/a$ 变为 $\Psi_{ab}(t)$ 时，其时窗半径 Δ_t 也相应地被压缩了 $1/a$. 这些都可以从几何角度得到形象的理解.

与之类似，关于 $\hat{\Psi}_{ab}(\omega)$ 的窗的度量也可用 $\hat{\Psi}(\omega)$ 的相应量来表示. 事实上，

因为由式(4)有

$$\hat{\Psi}_{ab}(\omega) = a^{-1/2} e^{-i\omega\frac{b}{a}} \hat{\Psi}\left(\frac{\omega}{a}\right),$$

$$|\hat{\Psi}_{ab}(\omega)|^2 = a^{-1}\left|\hat{\Psi}\left(\frac{\omega}{a}\right)\right|^2,$$

$$\|\hat{\Psi}_{ab}(\omega)\|_0^2 = \int_{\mathbf{R}} |\hat{\Psi}_{ab}(\omega)|^2 d\omega = \int_{\mathbf{R}}\left|\hat{\Psi}\left(\frac{\omega}{a}\right)\right|^2 d\left(\frac{\omega}{a}\right)$$

$$= \|\hat{\Psi}(\omega)\|_0^2, \tag{5}$$

于是有

$$\begin{aligned}\omega^* &= \frac{1}{\|\hat{\Psi}_{ab}(\omega)\|_0^2}\int_{\mathbf{R}} \omega\, |\hat{\Psi}_{ab}(\omega)|^2 d\omega \\ &= \frac{a}{\|\hat{\Psi}_{ab}(\omega)\|_0^2}\int_{\mathbf{R}}\left(\frac{\omega}{a}\right)\left|\hat{\Psi}\left(\frac{\omega}{a}\right)\right|^2 d\left(\frac{\omega}{a}\right) \\ &= \frac{a}{\|\hat{\Psi}(\omega)\|_0^2}\int_{\mathbf{R}} \xi\, |\hat{\Psi}(\xi)|^2 d\xi \\ &= a\omega_{\hat{\Psi}}^*,\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta_\omega &= \frac{1}{\|\hat{\Psi}_{ab}(\omega)\|_0}\left\{\int_{\mathbf{R}}(\omega-\omega^*)^2\, |\hat{\Psi}_{ab}(\omega)|^2 d\omega\right\}^{1/2} \\ &= \frac{1}{\|\hat{\Psi}_{ab}(\omega)\|_0}\left\{\int_{\mathbf{R}}(\omega-a\omega_{\hat{\Psi}}^*)^2\left|\hat{\Psi}\left(\frac{\omega}{a}\right)\right|^2 d\left(\frac{\omega}{a}\right)\right\}^{1/2} \\ &= a\Delta_{\hat{\Psi}}.\end{aligned} \tag{7}$$

从式(7)可知,当 $\Psi(t)$ 被压缩 $1/a$ 变为 $\Psi_{ab}(t)$ 时,其频窗半径被拉宽了 a 倍. 总之,用 $\Psi_{ab}(t)$ 作窗函数时,小波变换能在时—频局部范围内分析信号,此时,

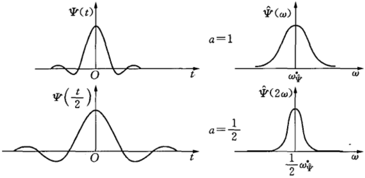
时—频窗中心为 $\left(\frac{1}{a}(t_{\Psi}^* + b),\ a\omega_{\hat{\Psi}}^*\right)$,

时窗宽度为 $2\Delta_t = 2\Delta_{\Psi}/a$,

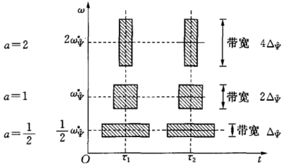
频窗宽度为 $2\Delta_\omega = 2a\Delta_{\hat{\Psi}}$,

时—频窗面积为 $2\Delta_t \cdot 2\Delta_{\hat{\omega}} = 4\Delta_{\Psi}\Delta_{\hat{\Psi}}$.

小波变换时—频窗如图 6.2.2 所示,其中图(a)表明,小波函数作为时窗函数,当它在时域被拉宽时,它的频窗宽度被压窄,它的频窗中心也变小;图(b)表明,对于同一个时窗中心而言,随着频窗中心(频带中心)上移,小波变换的频窗宽度(频带宽)被拉宽,相应地,其时窗宽度被压窄.



(a)



(b)

图 6.2.2　小波变换时—频窗示意图

下面具体说明小波变换的时—频局部化机理.

对于参数 a 固定、参数 b 自由的情形,小波变换 $W_f(a,\ b)$是关于变量 b 的时域函数;由于 $\hat{\Psi}_{ab}(\omega)$是频窗函数的缘故,小波变换 $W_f(a,\ b)$实际上是被限制在频窗 $[a\omega_{\hat{\Psi}}^* - a\Delta_{\hat{\Psi}},\ a\omega_{\hat{\Psi}}^* + a\Delta_{\hat{\Psi}}]$ 这个子频带范围内的时域函数.

对于参数 a 和参数 b 都固定的情形,由于 $\Psi_{ab}(t)$是时窗函数和 $\hat{\Psi}_{ab}(\omega)$是频窗函数的缘故,$W_f(a,\ b)$的时域和频域表现实际上被限制在时—频窗

$$[a\omega_{\hat{\Psi}}^* - a\Delta_{\hat{\Psi}},\ a\omega_{\hat{\Psi}}^* + a\Delta_{\hat{\Psi}}] \times \left[\frac{1}{a}(t_{\Psi}^* + b) - \frac{\Delta_{\Psi}}{a}, \frac{1}{a}(t_{\Psi}^* + b) + \frac{\Delta_{\Psi}}{a}\right]$$

范围内.因 $W_f(a,\ b)$是与 $f(t)$对应的一种积分变换,所以小波变换 $W_f(a,\ b)$实际上是在积分变换机制下将 $f(t)$和 $\hat{f}(\omega)$限制在时—频窗内的一种局部化表现.换句话说,$W_f(a,\ b)$ 在时窗内的表现对应着 $f(t)$ 在时窗内的表现,$FT[W_f(a,\ b)]$在频窗内的表现对应着 $\hat{f}(\omega)$在频窗内的表现.

下面进一步具体说明小波变换的时—频窗是自适应的.

从小波窗函数 $\Psi_{ab}(t)$ 的参数选择方面观察. 当 a 较大时, 频窗中心 $a\omega_{\hat{\Psi}}^{*}$ 自动地调整到较高的频率中心的位置, 且时—频窗形状自动地变为"瘦窄"状; 因为高频信号在很短的时域范围内的幅值变化大, 频率含量高, 所以这种"瘦窄"时—频窗正符合高频信号的局部时—频特性. 同样, 当 a 较小时, 频窗中心 $a\omega_{\hat{\Psi}}^{*}$ 自动地调整到较低位置, 且时—频窗的形状自动地变为"扁平"状; 因为低频信号在较宽的时域范围内仅有较低的频率含量, 所以这种"扁平"状的时—频窗正符合低频信号的局部时—频特性. WFT 仅具有不变时—频窗, 无论频率中心处于何处, 其时—频窗形状不改变, 其时—频局部分析显得很单一. 与之相比, 小波变换的时—频窗是灵活可调的, 是自适应的.

第四节　离散小波变换及其频带特性

在上一节中, 针对 a 和 b 是连续变量的情形, 分析了小波窗函数 $\Psi_{ab}(t)$ 关于时—频局部化功能和自适应的时—频窗的具体表现. 在实际应用中, 通常将 $\Psi_{ab}(t)$ 中的连续变量 a 和 b 取作整数离散形式, 将 $\Psi_{ab}(t)$ 表示为

$$\Psi_{j,k}(t) = 2^{j/2}\Psi(2^{j}t - k),$$

相应的小波变换表示为离散小波变换

$$W_f(j, k) = (f(t), \Psi_{j,k}(t))$$

的形式, 下面将进一步讨论离散小波 $\Psi_{j,k}(t)$ 和离散小波变换 $W_f(j, k)$ 所具有的子频带特性.

离散小波 $\Psi_{j,k}(t)$ 是由小波函数 $\Psi(t)$ 经 2^j 整数倍放、缩和经整数 k 平移所生成的函数族 $\{\Psi_{j,k}(t)\}$, $j, k \in \mathbf{Z}$, 它适合

$$\| \Psi_{j,k}(t) \|_0^2 = \| \Psi(t) \|_0^2,$$

$$| FT[\Psi_{j,k}(t)] |^2 = 2^{-j} | \hat{\Psi}(\omega/2^j) |^2.$$

若 $\Psi(t)$ 满足允许小波条件(本章第二节式(5)), 则 $\Psi_{j,k}(t)$ 也是允许小波; 若 $\Psi(t)$ 具有时—频局部化表现, 则 $\Psi_{j,k}(t)$ 也具有时—频局部化表现. 利用本章第三节式(6)和式(7)具体观察 $\Psi_{j,k}(t)$ 的频窗表现, 则

$$\hat{\Psi}_{j,k}(\omega) \text{的频窗中心为} \quad \omega^{*} = 2^{j}\omega_{\hat{\Psi}}^{*},$$

$$\hat{\Psi}_{j,k}(\omega) \text{的频窗半径为} \quad \Delta_{\omega} = 2^{j}\Delta_{\hat{\Psi}},$$

其中, $\omega_{\hat{\Psi}}^{*}$ 和 $\Delta_{\hat{\Psi}}$ 分别是 $\hat{\Psi}(\omega)$ 所具有的频窗中心和频窗宽度, 对选定的小波函数 $\Psi(t)$ 而言, 它们是确定量.

离散小波变换

$$W_f(j,\ k) = \int_{\mathbf{R}} f(t)\overline{\Psi}_{j,k}(t)\mathrm{d}t$$

是关于频率指标 j 和平移指标 k 的函数，由

$$(f(t), \Psi_{j,k}(t)) = \frac{1}{2\pi}(\hat{f}(\omega),\ \hat{\Psi}_{j,k}(\omega))$$

和 $\hat{\Psi}_{j,k}(\omega)$的频窗范围可知，经 $\Psi_{j,k}(t)$作用的小波变换实际上把信号$f(t)$的频率范围$\hat{f}(\omega)$限制在 $[\omega^* - \Delta_\omega,\ \omega^* + \Delta_\omega]$ 子频带内，小波变换结果是这个频带内的时域分量，小波$\Psi_{j,k}(t)$是一带通函数，它的小波变换在频域方面的局部化途径由 j 调节，在时域方面的局部化作用由 k 调节. 对于不同的 j，这些子频带是不重叠的，这样，原信号的频率范围被划分为由$\{\Psi_{j,k}(t)\}$所表现的子频带，信号经小波变换表现为不同子频带分量之和，对原信号的局部时—频分析就表现为对那些描述子频带内的时域分量的分析. 反应低频的局部分析应表现于反应低频的子频带分量中，反应高频的局部分析应表现于反应高频的子频带分量中，由此可知，小波分析方法并不像 Fourier 分析方法那样把时域信号表示为若干精确的频率分量之和，而是将其表示为若干描述子频带的时域分量之和. 正是由于这种表示方式，小波分析方法才获得了在时—频局部化方面的成功，才获得关于时—频局部化方面的丰富多彩的应用.

到目前为止，我们只是对小波变换的时—频局部化特点进行了初步介绍，为了深入研究和应用它，还必须努力构造小波函数.

此时，最容易想到的是 Gauss 函数，它在时域和频域中都有局部化特点，Morlet 小波函数

$$\Psi(t) = \mathrm{e}^{-t^2/2}\mathrm{e}^{\mathrm{i}\omega_0 t},$$

$$\hat{\Psi}(\omega) = \sqrt{2\pi}\mathrm{e}^{-(\omega-\omega_0)^2/2}$$

就是在这种思考下构造出来的. 另外，Marr 小波函数

$$\Psi(t) = (1 - t^2)\mathrm{e}^{-t^2/2},$$

$$\hat{\Psi}(\omega) = \sqrt{2\pi}\omega^2\mathrm{e}^{-\omega^2/2}$$

是利用 Gauss 函数 2 阶导数构造出来的. 这些小波函数都满足允许小波条件且具有时—频局部化特点. 还有，Shannon 小波和 Haar 小波也是容易想到和构造的. 为了不同的应用目的，还需构造出具有不同性能的小波函数，例如，为加强小波函数的光滑性、衰减性、紧支性、对称性以及正交性去构造各种小波. 历史上，人们为此作出了不少探索和尝试，但小波函数的构造方法还缺乏统一性.

1986 年，S. Mallat 和 Y. Meyer 提出了多分辨分析的框架，这是一个理性思考的又是具体可行的框架. 至此，我们才真正找到了构造各种小波函数的统一方法，也正是在此基础上，我们才在深层次上认识了小波函数和小波变换.

第三章　小波变换应用实例

本章主要讨论小波变换在信号处理与检测，语音信号处理中的应用；第一节主要讨论如何利用小波变换来进行信号的滤波，基于小波变换的时频受限信号的检测方法，瞬态信号的检测方法等内容，第二节主要讨论基于小波变换的语音编码方案以及基于小波变换的语音时域和频域基音提取等内容.

第一节　小波变换在信号处理与检测中的应用

一、小波分析用于信号滤波

设输人一维信号 $f(t)$，由于物理分辨率有限，可设 $f(x) \in V_{J_1}$，

$$f(x) = C_{J_1} f(x) = \sum_{k \in \mathbf{Z}} C_{J_1 k} \varphi_{J_1 k},$$

其中，$\{C_{J_1 k}\}_{k \in \mathbf{Z}}$ 具有有限长度 N_{J_1}，利用 Mallat 算法，我们将信号分解成不同的通道成分

$$f(x) = C_{J_2} f(x) + \sum_{j=J_2}^{J_1+1} D_j f(x),$$

其中，$C_{J_2} f(x) = \sum_{k \in \mathbf{Z}} C_{J_2 k} \varphi_{J_2 k}(x)$ 是信号频率低于 2^{-J_2} 的成分，而

$$D_j f(x) = \sum_{k \in \mathbf{Z}} D_{jk} \Psi_{jk}(x)$$

是 f 的频率介于 2^{-j} 与 $2^{-(j-1)}$ 的成分. 前面的小波分解式实际上可以写成

$$C_{j+1} = HC_j, \quad j = J_1, \cdots, J_2 - 1,$$

$$D_{j+1} = GC_j, \; j = J_1, \cdots, J_2 - 1,$$

其中，低通滤波器 H 作用在一个序列 $a = \{a_k\}_{k \in \mathbf{Z}}$ 的效果为

$$(Ha)_n = \sum_{k \in \mathbf{Z}} h_{k-2n} a_k,$$

高通滤波器 G 作用效果为

$$(Ga)_n = \sum_{k\in \mathbf{Z}} g_{k-2n} a_k,$$

其中，$\{h_k\}_{k\in \mathbf{Z}}$ 与 $\{g_k\}_{k\in \mathbf{Z}}$ 是由给定的多分辨率分析确定的镜像滤波器.

由上述分析可知，基于小波的多分辨率分析将信号 $f(x)$ 分解成频率小于 2^{-J_2} 成分 $C_{J_2}f(x)$ 和频率介于 2^{-j} 与 $2^{-(j-1)}$ 之间的成分 $D_j f(x)$，$(J_1+1\leqslant j\leqslant J_2)$. 如果尺度函数 φ 与小波函数 Ψ 的能量分别集中在 $x=a$ 和 $x=b$ 附近，则式中 $C_{jk}\varphi_{jk}(x)$ 和 $D_{jk}\Psi_{jk}(x)$ 能量分别集中在 $x=2^j(k+a)$ 和 $x=2^j(k+b)$ 附近.

按照 Mallat 算法将 $f(x)$ 分解之后，可以根据先验知识有效区分信号与噪声，加以滤波形成新序列 $\widetilde{C}_{J_2}$ 和 $\widetilde{D}_j(J_1+1\leqslant j\leqslant J_2)$，再按 Mallat 重建算法.

$$\widetilde{C}_{j-1} = H^* \widetilde{C}_j + G^* \widetilde{D}_j (j = J_2, J_2-1, \cdots, J_1+1),$$

得到去噪后的信号

$$\widetilde{f}(x) = C_{J_1}\widetilde{f}(x) = \sum_{k\in \mathbf{Z}} C_{J_1 k}\varphi_{J_1 k}(x).$$

二、基于小波变换的暂态信号检测方法

利用小波的伸缩和时移特性，通过对信号作多尺度的小波变换，可以在低信噪比下很好地检测到信号.

在很多实际应用中都会遇到暂态信号(transient signal)的检测问题，即在噪声中检测短时、非平稳、波形和到达时间未知的信号. 即考虑下面的检测模型：

$$H_0: x(t) = n(t),$$
$$H_1: x(t) = s(t) + n(t),\ t\in[0,\ T],$$

式中，$s(t)$ 只在 $[t_0,\ t_0+T_0]$ 内非零，$T_0 \ll T$，$n(t)$ 为零均值的 Gauss 白噪声，方差为 σ^2.

统计检测理论分析表明，已知信号波形时，最佳检测器为匹配滤波器(matched filter)，而当信号波形未知时，则通称采用广义似然比检测(GLRT)对未知信号做最大似然估计. 如果能够有效地利用关于信号的某些先验知识，就可以提高检测器性能. 利用小波变换，可以提取信号的波形信息，因此可以在较低信噪比情况下检测到暂态信号.

连续信号 $f(t)\in L^2(R)$ 的小波变换为

$$Wf(a,\ b) = \int_{-\infty}^{+\infty} a^{-\frac{1}{2}} h\left(\frac{t-b}{a}\right) f(t)\,\mathrm{d}t,$$

其中，a，b 为尺度因子和时移因子.

将 a，b 离散化，$a = a_0^m$，$b = na_0^m b_0$，($a_0 > 1$，m，$n \in \mathbf{Z}$)，得到的小波级数为

$$C(m,\ n) = \int h_{mn}(t) f(t) \mathrm{d}t,$$

式中

$$h_{mn}(t) = a_0^{-\frac{m}{2}} h(a_0^{-m} t - nb_0).$$

设暂态信号为

$$f(t) = \sum_{i=1}^{I} A_1 \exp(-a_1(t - t_1)) \sin(\omega_1(t - t_1) + \varphi_1) u(t - t_1),$$

其中，A_1，a_1，t_1，ω_1，φ_1 分别是第 i 个信号幅度、衰减系数、到达时间、频率和初相参数.

经过简单推导，就可以得到以下结论：

① $C(m,\ n)$在 $n = t_1 / b_0 a_0^m$ 两边呈指数衰减，而 $n = t_1 / b_0 a_0^m$ 处达到峰值，这个峰值的大小与 a_0，m，ω_1 等有关. 如果对信号作多个尺度的小波变换，在不同的尺度 a_0^m 上，与信号出现的时刻相对应，小波系数均会在$t_i / b_0 a_0^m$处出现峰值. 利用这一特点，在强噪声背景情况下，将一维信号变换到时间一尺度域，从而准确检测到信号.

② 小波变换是一种多分辨率分析，当 a_0^m 较小时，时间分辨率高，适合于分析高频信号，可以深入观察过程的细节，更加精确估计信号出现时间；当 a_0^m 较大时，频率分辨率较高，适合于分析低频信号，我们可以看到过程全貌. 这使得基于小波变换的检测器具有较好的鲁棒性.

③ 如果在观测时间内有多个信号，即 $I > 1$，这些信号可能在时间上部分重叠，此时可以选择 a_0^m，b_0，使$a_0^m b_0$尽可能小，从而使小波系数峰值出现在不同的 n 上，有效进行分离.

上述过程可以用离散化模型表示为

$$H_0:\ x = n,$$

$$H_1:\ x = s + n,$$

其中，x 为接收信号 M 维向量，s 为待检测信号 M 维向量，n 是协方差阵为 $\sigma^2 I$ 零均值 Gauss 白噪声 M 维向量. 小波变换可以看成是将信号向量乘一个各行是小波变换算子的采样值的变换矩阵 R，得到小波系数 $\{C(m,\ n) \mid M_1 \leqslant m \leqslant M_2,\ 0 \leqslant n \leqslant N-1\}$. 为了方便，仍用 x，s，n 等表示这些系数组成的向量.

在 H_0 下，x 是零均值、协方差阵为 Λ 的正态过程，因为小波变换是线性的，

$\Lambda = R\Delta R^{\mathrm{T}}$，如果噪声是白色的，就可以近似认为是对角阵. 而在 H_1 下，x 的所有$C(m, n)$中，仅与信号相对应的系数 $\{C(m, n) \mid M_1 \leqslant m \leqslant M_2, n \in S\}$ 的均值非零，S 为 $[0, N-1]$ 中 K 个整数的集合，其余的$C(m, n)$则与噪声相对应.

将 x 重排，使 x 的前 P 个元素为均值非零的元素，记为 x_1，后面 $L-P$ 个元素记为 x_2，相应的将 Λ 重排并分块为 $\begin{bmatrix}\Lambda_{11} & \Lambda_{12}\\ \Lambda_{21} & \Lambda_{22}\end{bmatrix}$，其中，$L = N(M_2 - M_1 + 1)$，$p = K(M_2 - M_1 + 1)$，则可以证明，广义似然比为

$$L_{WT}(X) = X^{\mathrm{T}}\Lambda^{-1}X - X_2{}^{\mathrm{T}}\Lambda_{22}^{-1}X_2.$$

在 H_0 下，$L_{WT}(x)$为 χ^2 分布，自由度为 P. 在 H_1 下，$L_{WT}(x)$为非中心 χ^2 分布，自由度为 P，非中心参数为 $m_1^{\mathrm{T}}(\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21})^{-1}m_1$. 其中 m_1 为 x_1 的均值.

将基于小波变换的检测器与下面两种检测器的性能做一比较：

① 匹配滤波器

$$L_{MF} = X_S^{\mathrm{T}}/\sigma^2,$$

L_{MF}在 H_0 下为零均值、方差为$S^{\mathrm{T}}S/\sigma^2$的正态分布，在 H_1 下为均值、方差都为$S_S^{\mathrm{T}}/\sigma^2$的正态分布.

② 能量检测器

$$L_{ED} = X^{\mathrm{T}}X/\sigma^2,$$

L_{ED}在 H_0 下是自由度为 M 的 χ^2 分布，在 H_1 下是自由度为 M 的非中心 χ^2 分布，非中心参数为$S^{\mathrm{T}}S/\sigma^2$.

在 Neyman-Pearson 准则意义下，给定某一虚警概率 P_f，我们来计算检测概率 P_d 与信噪比 SNR 的关系. 这里 SNR 定义为信号能量与噪声方差之比. 通过固定虚警概率 P_f，比较检测器的性能，可以发现基于小波变换的检测器的性能大大好于能量检测器，而与匹配滤波器相差不远. 基于小波变换的检测器利用单边指数小波，并通过调整小波的尺度和时移因子，小波变换系数较好地反映了信号的特征，从而能在较低信噪比下正确检测到信号和估计到达时间.

三、基于小波变换的信号最佳接收器的设计

通常，一个正交变换应该反映信号的特征，其中，KL 变换是近乎最佳的方案. 当信号的自相关矩阵在 KL 基中展开时，系数是互不相关的. 这种性质可以用来设计最佳信号检测结构. 经典的基于小波变换的方法是假设存在某个小波基作为输入信号的特征函数，信号检测结构就是寻找这样的特征小波，也就是在寻找一组变量的最佳值. 本节中将着重说明这些变量实际上是一个小波基用另

一个表示时的小波系数.本节采用的递推结构设计了基于小波变换的最佳接收器.这种结构对噪声中的信号检测是十分有益的.

通常的小波变换可以表示为

$$x(t) = \sum_k \sum_j c_{jk} \Psi_{jk},$$

其中

$$\Psi_{jk}(t) = 2^{\frac{j}{2}} \Psi(2^j t - k),$$

$$c_{jk} = \langle x(t), \Psi_{jk}(t) \rangle,$$

符号 $\langle x(t), \Psi_{jk}(t) \rangle$ 表示内积.

在正交变换中,假设基函数为 $\phi_k(t)$,对应的系数为 x_k,令 μ_k 表示 x_k 的均值.假定系数之间是不相关的,则

$$E[(x_k - \mu_k)(x_j - \mu_j)] = \lambda_k \delta_{kj}$$

成立的充要条件是

$$\lambda_k \phi_k(t) = \int K_x(t, u) \phi_k(u) \mathrm{d}k,$$

其中,$\phi_k(u)$为对应于特征值 λ_k 的特征函数,$K_x(t, u)$为自相关矩阵.

$$K_x(t, u) = E[(x(t) - \mu_x(t))(x(u) - \mu_x(u))].$$

如果 $K_x(t, u)$是正定的,则所有的特征值是正的.所有的特征函数构成了完备的正交基,这就是通常所说的 KL 变换.自相关矩阵 $K_x(t, u)$可以写成

$$K_x(t, u) = \sum_k \lambda_k \phi_k \phi_k(u).$$

当然也可以使用小波函数作为 KL 基.对于小波系数 c_{jk},设均值为

$$\mu_{kj} = E[c_{kj}] = \int \mu_x(t) \Psi_{jk}(t) \mathrm{d}t,$$

如果小波系数之间是不相关的,即

$$E[(c_{jk} - \mu_{jk})(c_{lm} - \mu_{lm})] = \lambda_{jk} \delta_{jl} \delta_{km},$$

则

$$E[(c_{jk} - \mu_{jk})^2] = \lambda \geqslant 0.$$

为了保证公式成立,则 $\Psi_{jk}(t)$必须为输入信号的自相关矩阵的特征函数

$$\lambda_{jk} \Psi_{jk}(t) = \int K_x(t, u) \Psi_{jk}(u) \mathrm{d}u.$$

在最佳接收器中，使用均方误差准则来预测迭加了白噪声 $\eta(t)$ 的信号 $\alpha(t)$. 在有色噪声情况下，可以对有色噪声进行预测，然后从接收信号中减去它. 设输入信号 $r(t)=\alpha(t)+\eta(t)$ 的自相关矩阵可以用小波基表示为

$$K_\alpha(t,\ u)=\sum_{jk}\lambda_{jk}\Psi_{jk}(t)\Psi_{jk}(u),$$

则最佳接收器的响应 $\tilde{\alpha}(t)$ 应该能够最小化 $\alpha(t)$ 与 $\tilde{\alpha}(t)$ 均方误差距离. 它可以由线性滤波器

$$h_0(t,\ u)=\sum_{jk}\frac{\lambda_{jk}}{\lambda_{jk}+N_0/2}\Psi_{jk}(t)\Psi_{jk}(u)$$

的脉冲响应得到

$$\tilde{\alpha}(t)=\sum_{jk}c_{jk}h_{jk}\Psi_{jk}(t),$$

这里 c_{jk} 为输入信号 $r(t)$ 的小波变换系数. 而

$$h_{jk}=\frac{\lambda_{jk}}{\lambda_{jk}+N_0/2}.$$

但是特征值和特征函数很难得到，给实际检测带来困难. 为此，我们考虑另外一个不是特征函数的小波函数 $\varphi_{jk}(t)$. 令 a_{pq}^{jk}, d_{pq} 分别表示 $\Psi_{jk}(t)$ 和输入信号 $r(t)$ 在小波函数 $\varphi_{pq}(t)$ 中的展开系数. 则有

$$\tilde{\alpha}(t)=\sum_{rs}d_{rs}\sum_{pq}\varphi_{pq}(t)b(p,\ q,\ r,\ s).$$

这里

$$b(p,\ q,\ r,\ s)=\sum_{jk}h_{jk}a_{pd}^{jk}a_{rs}^{jk}.$$

可以使用自适应过程的最小均方误差(LMS)算法来设计滤波器，系数 $b(p,\ q,\ r,\ s)$ 随着输入信号变化来最小化均方误差. 在自适应过程中，是不需要关于特征函数的先验知识的. 相反，先验知识可以简化自适应滤波器的设计过程.

如果考虑的时间范围为 $[0,\ T]$，设计最佳接收滤波器的问题就变成寻找最小化 $E\left[\int_0^T(\alpha(t)-\tilde{\alpha}(t))^2\mathrm{d}t\right]$ 的变量 h_{jk}, a_{pq}^{jk}, a_{rs}^{jk} 的问题. 可以注意到这些变量之间不是相互独立的. 假定输入 $r(t)$ 和信号 $\alpha(t)$ 对小波基 φ, Ψ 而言都局限于尺度 0 和 $J>0$ 之间，并且在所考虑的时间窗内平移参数都局限于 $[-K,\ K]$ 之间 (K 为整数)，则有

$$r(t)=\sum_{j=0}^{J}\sum_{k=-K}^{K}d_{jk}\varphi_{jk}(t)=\sum_{j=0}^{J}\sum_{k=-K}^{K}c_{jk}\Psi_{jk}(t).$$

$\varphi_{jk}(t)$和特征函数 $\Psi_{jk}(t)$之间可以互相表示为

$$\varphi_{jk}(t) = \sum_{pq} a_{jk}^{pq} \Psi_{pq}(t),$$

$$\Psi_{jk}(t) = \sum_{pq} a_{pq}^{jk} \varphi_{jq}(t),$$

其中

$$a_{pq}^{jk} = \int \Psi_{jk}(t) \varphi_{pq}(t) \mathrm{d}t.$$

同样地有

$$c_{jk} = \sum_{pq} a_{pq}^{jk} d_{pq},$$

$$d_{pq} = \sum_{jk} a_{pq}^{jk} c_{jk},$$

$$\sum_{pq} a_{pq}^{jk} a_{pq}^{lm} = \sum_{pq} a_{jk}^{pq} a_{lm}^{pq} = \delta((j, k), (l, m)). \tag{1}$$

由式(1)可以推出

$$a_{jk}^{pq} = \int 2^{\frac{j}{2}} \Psi(2^j t - k) 2^{\frac{p}{2}} \varphi(2^p t - q) \mathrm{d}t.$$

上式稍作变化就可以写为

$$a_{pq}^{jk} = \int \Psi(t) 2^{\frac{p-1}{2}} \varphi(2^{p-j} t - (q - 2^{p-j} k)) \mathrm{d}t. \tag{2}$$

或

$$a_{pq}^{jk} = \int 2^{\frac{1-p}{2}} \Psi(2^{j-p} t - (k - 2^{j-p} q)) \varphi(t) \mathrm{d}t. \tag{3}$$

由式(2)和式(3)可以得到递推过程如下：

$$a_{pq}^{jk} = a_{p-j,\ q-2^{p-j}k}^{00},\ p \geqslant j,$$

$$a_{pq}^{jk} = a_{00}^{j-p,\ k-2^{j-p}q},\ j \geqslant p.$$

这个递推过程可以用来简化像式(1)中 $c_{jk} = \sum_{pq} d_{pq} a_{pq}^{jk}$ 这样的项的计算. 为了讨论最佳检测的结构，从尺度因子 $j = 0$ 开始计算，它用矩阵形式可以表示为

$$c_{jk} = \mathrm{Tr}\{DA_{jk}\},$$

这里 D 和 A_{jk} 分别是 $(J+1)\times(2K+1)$ 和 $(2K+1)\times(J+1)$ 的矩阵，D 定义为

$$D=\begin{bmatrix} d_{0,-K} & d_{0,-K+1} & \cdots & d_{0,K-1} & d_{0,K} \\ d_{1,-K} & d_{1,-K+1} & \cdots & d_{1,K-1} & d_{1,K} \\ \vdots & \vdots & & \vdots & \vdots \\ d_{J-1,-K} & d_{J-1,-K+1} & \cdots & d_{J-1,K-1} & d_{J-1,K} \\ d_{J,-K} & d_{J,-K+1} & \cdots & d_{J,K-1} & d_{J,K} \end{bmatrix}, \tag{4}$$

A_{0k}定义为

$$A_{0k}=\begin{bmatrix} a_{0,-K-k} & a_{1,-K-2k} & \cdots & a_{J,-K-2^Jk} \\ a_{0,-K+1-k} & a_{1,-K+1-2k} & \cdots & a_{J,-K+1-2^Jk} \\ \vdots & \vdots & & \vdots \\ a_{0,K-1-k} & a_{1,K-1-2k} & \cdots & a_{J,K-1-2^Jk} \\ a_{0,K-k} & a_{1,K-k-2k} & \cdots & a_{J,K-2^Jk} \end{bmatrix}.$$

这里,矩阵中元素省略了上标“00”,可以从式(4)看出,对于 $j>0$ 的更一般情况,矩阵中元素是序列 a^{00} 和 a_{00} 的组合形式. 它们之间满足式(2)关系.

四、信号瞬时特征的小波分析方法

本节主要讨论如何利用平稳相位逼近技术获取渐近小波变换的性质,从小波系数的辐角参数中提取小波脊线,再由小波脊线提取信号的瞬时特征量. 为了计算小波脊线,还引入了小波曲线.

设瞬态信号为

$$s(t)=A_s(t)\cos\varphi_s(t),$$

其中

$$\varphi_s'(t)\gg A_s'(t),$$

信号 $s(t)$的瞬时频率为

$$v_s(t)=\frac{1}{2\pi}\frac{\mathrm{d}\varphi_s(t)}{\mathrm{d}t}.$$

信号分析和目标特征提取的往往是信号的幅度 $A_s(t)$和频率 $v_s(t)$等瞬时特征量. 对多个渐近信号的合成信号,还需要分离和提取出信号的各分量的瞬时特征. 利用小波分析方法可以准确地提取出被分析信号的幅频调制特性.

设 $g, h\in H^2(R)$,它们满足

$$0<|C_{gh}|<\left|\frac{\int_0^\infty g^*(\omega)h(\omega)\mathrm{d}\omega}{\omega}\right|<+\infty,$$

其中 $h(\omega)$, $g(\omega)$为 h, g 的 Fourier 变换, $H^2(R)=\{f\mid f\in L^2(R),\ f(\omega)=0,$

当 $\omega \leqslant 0$ 时}为实 Hardy 空间. g 叫做分析小波,h 叫做重构小波. 记由伸缩和平移构成的小波基为

$$\overset{\circ}{g_{ba}}(t) = \frac{1}{a} g\left(\frac{t-b}{a}\right),\ h_{ba}(t) = \frac{1}{2} h\left(\frac{t-b}{a}\right).$$

记小波变换为

$$T_s(b,\ a) = \langle s,\ g_{ba} \rangle = \int_{-\infty}^{+\infty} s(t) g_{ba}^*(t) \mathrm{d}t,$$

$T_s(b,\ a)$叫做小波系数,则信号由小波系数得到的时间尺度表征为

$$s = \frac{1}{C_{gh}} \cdot \frac{\int_0^{\infty} \int_{-\infty}^{+\infty} T_s(b,\ a) h_{ba} \mathrm{d}a\mathrm{d}b}{a}.$$

小波系数可以用幅度和相位表示为 $T_s(b,\ a) = \{|\ T_s(b,\ a)\ |,\ \arg[T_s(b,\ a)]\}$,利用小波系数的幅角信息 Ψ_{ba},就可以提取信号的瞬时特征.

设一渐近单频信号为 $s(t) = A(t)\cos\varphi(t)$,利用 Hilbert 变换得到其解析信号为

$$Z_s(t) = (1 + jH)s(t) = A_s(t)\exp\{j\varphi_s(y)\}.$$

假定基本小波 $g(t)$本身是渐近的,可表示为

$$g(t) = A_g(t)\exp\{j\varphi_g(t)\},$$

则有

$$\langle s,\ g_{ba} \rangle = \langle Z_s,\ g_{ba} \rangle.$$

可得到信号的渐近小波系数为

$$T_s(b,\ a) = \langle Z_s, g_{ba} \rangle = \frac{1}{2a} \int_{-\infty}^{+\infty} M_{ba}(t)\exp\{j\varphi_{ba}(t)\}\mathrm{d}t,$$

$$M_{ba}(t) = A_s(t) A_g^*\left(\frac{t-b}{a}\right),$$

$$\varphi_{ba}(t) = \varphi_s(t) - \varphi_g\left(\frac{t-b}{a}\right).$$

假设对任意的$(b,\ a)$, $\varphi_{ba}(t)$关于 t 只有一个一阶驻点 $t_s = t_s(b,\ a)$, 并且 $\varphi'(b,\ a) = 0$, $\varphi''(b,\ a) \neq 0$, 则$T_s(b,\ a)$的渐近展开可以记为

$$T_s(b,\ a) = \frac{Z_s(t_s) g^* \left(\frac{t_s - b}{a}\right)}{\mathrm{Corr}(b,\ a)},$$

$$\mathrm{Corr}(b,\ a) = \overline{2/\pi a^2\ |\ \varphi''_{ba}(t_s)\ |}\exp\{-j\pi/4\operatorname{sgn}[\varphi''_{ba}(t_s)]\}.$$

为了分析渐近解析信号的小波变换的特征，下面引入小波脊线和小波曲线的概念.

小波脊线定义为

$$R = \{(b, a) \in H^2(R), t_s(b, a) = b\}.$$

在小波脊线上有

$$\varphi'_{ba}(t_s)\mid_b = 0,$$

$$\varphi'_g(0)/\varphi'_s(0) = a_r(b).$$

因此小波脊线可以用 $a_r(b)$ 来表示，它代表了信号的瞬时频率.

小波脊线上的那部分小波系数为

$$T_s(b, a_r(b)) = \{g(0)/\mathrm{Corr}[b, a_r(b)]\}Z_s(b).$$

当小波脊线 $a_r(b)$ 提取出来后，从小波系数 $T_s[b, a_r(b)]$ 可得到原始信号，从而得到信号振幅. 这样就可以得到信号的频率和振幅，分别为

$$v_s(b) = \frac{1}{2\pi}\varphi_s(b) = \frac{1}{2\pi}\frac{\varphi_g(0)}{a_r(b)},$$

$$A_s(b) = Z_s(b) = | T_s(b, a_r(b)\mathrm{Corr}[b, a_r(b)]/g^*(0) |.$$

定义小波曲线为

$$C = \{(b, a) \in H^2(R), t_s(b, a) = b_0\}.$$

它指在给定时刻 $t_s(b, a) = b_0$ 时，通过点 $(b_0, a_r(b_0))$ 的曲线.

在小波曲线上，有

$$T_s(b, a) = \frac{g^*\left(\dfrac{b_0 - b}{a}\right)Z_s(b_0)}{\mathrm{Corr}(b, a)},$$

令 $\Psi(b, a) = \arg\{T_s(b, a)\}$，则在给定的小波曲线上，$\Psi(b, a)$ 满足

$$\left.\frac{\partial\Psi(b, a)}{\partial b}\right|_{t_s(b, a)=b_0} = \frac{1}{a}\varphi'_g\left(\frac{t_s - b}{a}\right) + \frac{\partial a}{\partial b}\frac{t_s - b}{a^2}\varphi'_g\left(\frac{t_s - b}{a}\right)$$

$$= \left.\frac{\varphi'_g(0)}{a}\right|_{t_s=b} = \frac{\omega_0}{a},$$

即在给定的小波曲线上的瞬时频率等于小波脊线上的基函数的中心频率. 因此可以得到小波脊线提取的一般方法.

对采样周期为 T_s 的离散序列 $s_k = s(t_k)$，$k = 0, 1, \cdots, N-1$，记它在给定的伸缩因子 a 上的 Fourier 变换

$$T_a(k) = T_s(t_k, a) = \{ | T(t_k, a) |, \Psi_a(t_k) \},$$

D_b 为对平移因子 b 的离散微分算子，则有

$$D_b \Psi_a(t_k) = \frac{\omega_0}{a},$$

因此

$$a = \frac{\omega_0}{D_b \Psi_a(t_k)}.$$

令 $q_r^{(0)}(t_k)$ 为 $a_r(t_k)$ 的初始估计量，则可以通过下面的迭代方法来计算小波脊线

$$a_r^{(i+1)}(t_k) = \omega_0 / D_b \Psi a_r^{(i)}(t_k),$$

$$a_r^{(0)}(t_{k+1}) = a_r^{(i+1)}(t_k), t_k = t_0 + kT_0,\ k = 0, 1, \cdots, N-1.$$

所要求出的小波脊线为

$$a_r(t_k) = a_r^{(i+1)}(t_k).$$

有人曾经选用 Morlet 分析小波 $g(t) = \exp\{-t^2/2\}\exp\{j\omega_0 t\}$，对渐近信号进行了分析并作了计算机模拟. 结果表明，从渐近信号的连续小波变换的相位信息中提取小波脊线，可以快速准确地得到信号的瞬时特征，计算简单而且准确.

第二节　小波变换在语音信号处理中的应用

一、基于小波变换和矢量量化的语音编码

本小节主要讨论小波变换和矢量量化(vector quantization)的基本原理，并把矢量量化应用到小波变换域，从而实现一种低编码率、高信噪比的编码方案. 为了更清楚地帮助读者理解语音的小波分解过程，我们作小波分解过程示意图(图 6.3.1).

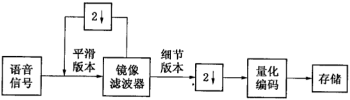


图 6.3.1　语音小波分解示意图

这里从矢量量化的最经典算法 LBG 算法谈起，自从 LBG 算法提出以来，矢量量化广泛应用于语音和图像编码中. 矢量量化应用的两个过程是码书设计和

码书搜索.设计高质量的码书可以改善矢量量化的性能,利用快速搜索可以使其实用化.我们发现小波变换域系数适合于矢量量化,编码模型如图6.3.2所示.

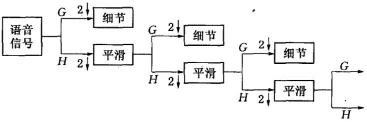


图 6.3.2 基于小波分析和矢量量化的编码模型

矢量量化可定义为 K 维欧氏空间 $\mathbf{R}^K$ 到 $\mathbf{R}^K$ 有限子集的映射 $\sigma: \mathbf{R}^K \to Y$,式中 $Y = \{X_i, i = 1, 2, \cdots, N\}$ 是所有码矢量的集合,称为码书,N 为码书大小,矢量量化可以分为两个映射:编码器和译码器,编码器将输入矢量 X_i 映射成码矢量的地址或标号 i,译码器将标号映射成码矢量,基本框图见图 6.3.3.

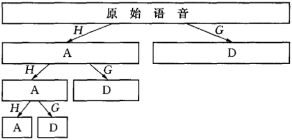


图 6.3.3 矢量量化基本图

为了降低码书搜索复杂度,采用二叉树结构的码书搜索和设计,从而大大降低了编码运算复杂度.

每次小波变换把信号分解成不同通道的成分即平滑版本和细节版本.每次分解相当于原信号减抽样成两个信号,而对这两个信号可采用比原信号更为有效的编码.对小波变换域系数的概率分布的研究表明,细节版本系数的典型分布为 Gauss 分布.对第 j 层分解得到的细节,其分布可表示为

$$p_j(x) = a_j \exp\{-|b_j x|^{r_j}\},$$

$$a_j = b_j r_j \Big/ 2\Gamma\left(\frac{1}{r_j}\right),$$

$$b_j = \frac{1}{\sigma_j}\Gamma\left(\frac{3}{r_j}\right) \Big/ \Gamma\left(\frac{1}{r_j}\right),$$

式中,r_j 为对应的参数,σ_j 为对应的方差.对其系数进行矢量量化,则量化增益可以表示为

$$G_{vq} = \frac{2^{-c}\left[\int p_j(x)^{\frac{1}{c+1}}\mathrm{d}x\right]^{c+1}}{(1+c)A(k,c)\left[\int p_j(x)^{\frac{1}{c+k}}\mathrm{d}x\right]^{c+k}}.$$

可见,G_{vq} 与 r_j 无关,K 为矢量量化维数.我们常用均方误差准则($c=2$),$A(k,2)$实际上表示了编码误差的上界.上式实际上是采用矢量量化理论上的最小增益.运用小波分析,得到原始语音的平滑版本和不同尺度上的一系列细节版本.对不同层上的细节信息进行矢量量化,应用 LBG 算法来设计矢量量化码书.我们对典型语音信号进行分析和训练以产生码书.训练序列由典型语音在对应分解层上的细节版本的一维矢量构成.初始码字由训练序列的形心(centriod)分裂得到.

整个码书由若干子码书构成,对应每层小波分解得到的细节版本分别有不同的子码书.编码时,只在对应的子码书中搜索,而不是在整个码书中搜索.

考虑到不同分解层上的细节版本重要性不同,我们对它们采用不同的参数进行矢量量化码书设计.显然,从更平滑版本中经小波分解得到的细节版本更加重要一些,随着 j 的增加第 j 层细节版本对恢复语音信号质量的影响程度是增加的.因而允许的量化失真必须更小,反映在码书设计中即采用较小的量化维数、较大尺寸的码书和较为严格的收敛条件,由此得到码书的构成如图 6.3.4.

$j=1$ 层细节版本子码书
$j=2$ 层细节版本子码书
…… ……
$j=J$ 层细节版本子码书

图 6.3.4　细节版本码书构成示意图

实际上,平滑版本对应变化缓慢的低频信息,对语音质量的影响较大,而细节版本对应高频信息,对应人耳不太敏感的频率.这是选择量化参数的直接依据.

对平滑版本,可采用矢量量化,也可采用其他量化方法.

下面估算编码压缩比及运算复杂度.用上述方案时,总的压缩比可表示为

$$C = \log_2 I/R_i,$$

$$R_i = R_J^{SQ}/2^{2j} + \sum_j R_j/2^{2j},$$

$$R_j = \log_2 V_j/K_j,$$

其中，K_j 为第 j 层细节版本矢量量化维数，V_j 为对应码书大小，而 R_i^{SQ} 为平滑版本编码所需要的比特数.

当采用长度为 R 的有限长度滤波器时，小波分解复杂度同前述分析. 当采用二叉树搜索方式时，设码书大小为 V，每量化一个矢量需 $\log_2 V$ 次. 设第 j 层有 L_j 个矢量，量化维数 K_j，则需要 $O(\sum_j L_j K_j \log_2 V_j)$ 次操作，与全搜索方式相比较，后者需要 $O(\sum_j L_j K_j V_j)$ 次，可见运算复杂度大大降低，利于实时实现.

我们在小波变换域进行矢量量化，而不是直接进行矢量量化，因为小波变换将原始信号 f 分解成不同频率通道的划分. 又把每通道成分按相位进行分解——频高者密，频低者疏. 这些特点使得我们能在较高编码率情况下，仍然有较好的重建语音质量. 可以断言，对于常规的一维信号，都可以应用这种思想.

二、基于小波变换的基音提取

我们结合小波变换与多分辨率分析，在频域合理引入通过峰值检测提取基音周期的方法. 由于这样处理后展开了原始信号的低频段的频域波形，不仅大大减小了整数倍基音周期的误判概率，而且提高了基音周期估计的精度. 整个算法速度快，适合于数字信号处理器实现.

由离散小波变换可知，对一个信号序列进行小波变换，可以在时域上得到原始信号的越来越低分辨率的版本，采样率以 $\frac{1}{2}$ 速率递减，在信号的时—频平面上，根据Heissenberg不等式

$$\Delta t \Delta f \geqslant 1/4\pi$$

可以看出，时域分辨率的降低意味着频域分辨率的相应提高，其中 Δt 为时域分辨率，Δf 为频域分辨率.

从镜像滤波器的角度来看，小波分析是以两组正交的高低通滤波器对输入信号进行滤波，得到高频和低频分量，并进行 $\frac{1}{2}$ 减抽样. 这样，对低频分量而言，时域分辨率降低一倍，而频域分辨率则相应提高一倍.

为了验证这种思想，我们构造了一个包含 200 Hz，300 Hz，400 Hz，531 Hz，670 Hz，750 Hz，800 Hz 的等频率分量的多频混合信号，并以 8 kHz 速率采样，得到 $s(n)$，分析 $s(n)$的 FFT 谱.

我们先对 $s(n)$ 作两次小波变换，得到两次低通滤波并 $\frac{1}{4}$ 减抽样的序列 $\hat{s}(n)$，再进行 FFT 分析其幅度谱. 实验结果表明，经两次小波变换后，频域低段分辨率得以提高，将原来混迭的频谱展开了 4 倍. 频谱峰值明显展开. 如果我们

对变换后的频谱曲线估计峰值位置，显然可以得到较为精确的结果.

对上述 $s(n)$进行处理，得到检测结果见表 6.3.1. 可以看出，在频域上能用 FFT 整数抽样表示的频率附近，变换前后均可得到较好的结果. 但在远离整数抽样点表示的频率处，变换后估计结果明确优于变换前的估计结果. 在语音基音周期估计中引用此结果，可以提高语音基音估计精度.

表 6.3.1　小波变换前后频率估计结果(频率单位:Hz)

	小波变换前			小波变换后			小波变换-内插		
实际频率	峰值位置	估计频率	误差	峰值位置	估计频率	误差	峰值位置	估计频率	误差
200	6	187	13	26	203	3	26	203	3
300	10	312	12	38	297	3	38	297	3
531	17	531	0	68	531	0	68	531	0
670	21	656	14	86	672	2	86	672	2
750	24	750	0	96	750	0	96	750	0
800	26	812	12	102	797	3	104	812	12

在语音信号基音估计中引入小波变换的主要障碍为帧长效应. 由于语音分帧处理时，语音为准平稳信号，不能取很大的帧长，在实时处理时考虑到时延更是如此. 对每帧有限点语音进行处理时经小波变换和减抽样，只能得到比原来较少的数据量，这时在其后补零，扩展为原始帧长再进行 FFT. 事实上，这种做法仍然能够保留理想情况. 为证明此点，我们将波形限于 256 点，作两次小波变换得到 64 点数据，为减小信号突变引入的高频分量，先对其加 Hamming 窗，再扩展为 256 点，分析其 FFT 幅度谱. 我们发现，这样处理后的信号幅度谱仍然保留了理想小波变换后幅度变化趋势，特别是峰值位置，为实际的基音估计应用小波变换奠定了基础.

因两次小波变换将基音峰值位置展开了 4 倍，我们至少可以获得$\frac{1}{4}$点估计精度. 若是检测更后面的峰值位置可以满足$\frac{1}{8}$，$\frac{1}{16}$点精度要求.

可以利用小波变换的带通性质和多分辨率性质，对语音进行逐次低通处理，滤除其高频谐波，只保留其低频成分，进而提出从时域在这样的分辨率的平滑版本下估计基音周期的方法. 由于这种方法滤除了语音信号的高次谐波，可以提高估计精确度.

从滤波器角度来看，小波变换实际上就是把信号分成高低频分量，并进行$\frac{1}{2}$

减抽样.这种变换实际上是对输入信号用小波基卷积后减抽样的结果.算法简单,利用小波变换的滤波性质,对原始语音信号进行小波变换,但不进行抽样,对保留的成分再类似逐级处理,可以得到原始语音的逐次平滑版本.

为此,我们在实际处理中对一段语音进行 4 次小波变换,观察其时域与频域波形,可以看出,随着变换层数的增加,原始语音中高次谐波成分逐渐减少,到第四层平滑版本已经几乎只有基频成分了.我们从这样的平滑版本统计时域峰值点各数进行基音周期估计,可以得到较为精确的结果.

在实际处理中发观,4 次小波变换不能有效清除二次谐波成分,我们曾试着继续分解下去,发现滤除效果改善不大,相反却大大降低了语音基频幅度.

小波变换虽然未能完全滤除二次谐波成分,却大大降低了二次谐波幅度.在第 4 次小波变换结果中,正电平峰值与负电平谷值信号至少一侧体现了基音周期信息.第四层小波变换时域波形峰值体现了本帧语音基音信息.为此,在估计基音周期时,同时统计正电平峰值和负电平谷值信息,以较大的时域基音估计值为有效估计结果.这样处理后,能从第四层小波变换时域波形准确估计基音周期.

实际处理中,为避免误提,第一次求出的基音周期用来修正估计结果,剔除偏差较大的峰值点或谷值点,再进行一次估计基音周期,大大提高估计精确度,这样求出的时域基音周期有很高精度.但由于使用整数点表示,稍有不足.为此,以时域基音周期为初值,变换到频域中,用频域峰值法在基音估计值附近进行调整,可使估计精度满足要求.由于分时域粗估和频域细估两部分,大大减少了频域基音搜索范围,提高了频域峰值估计速度.

三、基于小波变换的语音特征参数的提取

考虑到语音信号的时频信息,可以将语音信号作为非平稳信号来研究,在小波变换的基础上提出一种新的语音参数——WT 参数,它能够满足人耳听觉系统对声音信号常 Q 感知的特性.基于语音信号短时平稳的线性预测系数(LPC)特征提取方法,把语音产生的模型的参数提取出来作为特征,在一定程度上抓住了语音信号的特点,取得了一定的成功.

生理学的研究表明人类的听觉特性是一个常 Q 感知系统,这就启发我们去寻找一种既具有常 Q 特性,又可去掉短平稳假设的分析方法来分析语音信号,并以此作为特征参数.小波变换为这一思想提供了理论上的可能性.对于语音信号,必须用时变非平稳的分析方法才能很好地描述.

与 LPC 参数不同的是,这里提出的 WT 参数不是模型化参数.在一定的条件下,可以选取不同类的小波来计算此参数,这种参数去掉了短时平稳的假设,并且反应了语音的常 Q 特性.

设$\{V_j\}$是一给定的多分辨性分析，φ与Ψ分别为相应的尺度函数与小波函数，$f(x)$是属于V_1的任一信号，可以由 Mallat 算法进行分解. 经过分解，得到语音信号在不同尺度下的平滑版本和细节版本(小波表示)，其中小波表示就是我们的特征参数.

因为语音信号频带不越过 20 kHz，因此利用 Mallat 算法时，我们用第五级的小波表示为特征即可，因此得到 WT 参数的计算流程如图 6.3.5 所示.

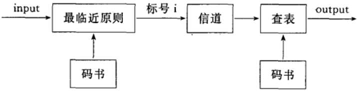


图 6.3.5　WT 参数计算流程

有人曾经对 WT 参数的提取做了一些工作. 他们把从话筒接收的语音信号首先用截止频率 5 kHz 的低通滤波器进行滤波，然后用采样频率为 10 kHz，精度 12 bit 的 A/D 变换器进行采样，得到语音信号，采用 Mallat 算法，以第五层小波表示作为 WT 参数. 从实验结果来看，很大程度上抑制了个人信息，而且对噪声不敏感，鲁棒性强，对特别是爆破音这类非平稳性不好的信息有很好的描述.

参 考 文 献

1. 秦前清,等.实用小波分析.西安:西安电子科技大学出版社,1993.
2. 刘贵忠,邸双亮.小波分析及其应用.西安:西安电子科技大学出版社,1992.
3. 王仁华等.语音信号处理.合肥:中国科学技术大学出版社,1998.
4. 孙功宪.多媒体数据库中的压缩编码技术.合肥:中国科学技术大学出版社,1994.
5. 徐佩霞,孙功宪.小波分析与应用实例.合肥:中国科学技术大学出版社,1996.

第七篇　分形理论及其应用

分形理论是美籍数学家 Mandelbrot(曼德尔布洛特)于 20 世纪 60—70 年代提出和建立的,现已成为处理自然科学及工程中许多不规则图形的强有力工具. 分形(fractal)一词,是 Mandelbrot 在 1975 年首次提出的,其原意是不规则、支离破碎等意义,而分形理论主要研究不规则、处处连续而处处不可微分的具有长程相关性、非平衡性的统计过程,借助于事物内部的自相似性质来洞察隐藏于混沌现象中的精细结构. 作为一门新兴的非线性技术,分形理论不断冲击着自然科学、社会科学和思维科学的各个领域,在各个领域都得到了广泛的应用. 分形理论为不同学科发现新的规律提供了崭新的语言和定量的描述,为现代科学的深入发展提供了新的思路和方法. 本篇介绍分形理论的初步知识和应用.

第一章　分形的例子

长期以来,人们习惯在 Euclid 空间 $\mathbf{E}^n$ 中对事物进行度量,并且建立了 Euclid 几何来对事物进行刻画. 但 Euclid 几何的要素是直线,平面,圆等,而现实生活中"云不是球,山岳不是锥体,海岸线不是圆,树皮不是光滑的,闪电也不是沿直线传播"(Mandelbrot). 显然 Euclid 几何不足以描述现实的复杂性,分形几何为这类复杂性的描述提供了新的思想和方法.

第一节　经典分形图形

一、Cantor 集

1883 年 Cantor 构造出如下的集合. 选取线段 $A_0 = [0, 1]$, 将其三等分并去掉中间的一段,记所得的集合为 $A_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$, A_1 的长度为 2/3. 对这两条较短的线段 $\left[0, \frac{1}{3}\right]$ 和 $\left[\frac{2}{3}, 1\right]$ 分别重复刚才的步骤,则得到集合 $A_2 =$

$\left[0,\ \frac{1}{9}\right] \cup \left[\frac{2}{9},\ \frac{1}{3}\right] \cup \left[\frac{2}{3},\ \frac{7}{9}\right] \cup \left[\frac{8}{9},\ 1\right]$，$A_2$ 的长度为 4/9. 按照这种方法继续下去得到 A_n（图 7.1.1），其长度为 $\left(\frac{2}{3}\right)^n$. 当 $n \to \infty$ 时就得到了 Cantor 集 A，显然 Cantor 集 A 是一个有限区间上的无穷集合，然而它的长度为 0.



图 7.1.1　Cantor 集

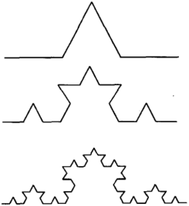


图 7.1.2　Koch 曲线

二、Koch(科赫)曲线

1904 年 Koch 设计了 Koch 曲线. 设 $A_0 = [0,\ 1]$，将其三等分并以中间一段为底作一个等边三角形，然后去掉中间的一段，得到一条折线 A_1，A_1 的长度为 4/3. A_1 由四条短直线段构成，对每一条短直线段重复刚才的步骤，得到集合 A_2，A_2 的长度为 16/9. 按照这种方法继续下去得到 A_n（图 7.1.2），其长度为 $\left(\frac{4}{3}\right)^n$. 当 $n \to \infty$ 时就得到了 Koch 曲线 A，显然 Koch 曲线 A 的长度为无穷. 在 Koch 曲线上的任意一点都不能作出 Koch 曲线的切线，即 Koch 曲线是处处连续但处处不可导的.

如果对一个等边三角形的各条边采用上述方法处理就得到一 Koch 雪花（图 7.1.3）. 实际上一朵 Koch 雪花由三条 Koch 曲线组成. 它是一条连续的闭曲线，不会自我相交且其长度是无限的. 但是 Koch 雪花的面积是有限的，不会超过初始三角形的外接圆的面积. 在有限的面积内可以画出长度无限的不会自我相交的曲线，这是一个与通常的“合理”直觉矛盾的结果.

三、Sierpinski(希尔宾斯基)垫

1915 年 Sierpinski 创造了一个面积为零，边长为无限的图形（图 7.1.4），人们称之为 Sierpinski 垫.

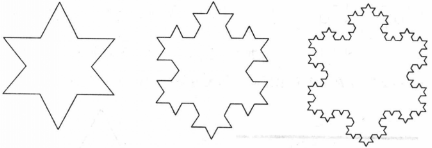


图 7.1.3　Koch 雪花

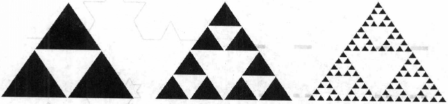


图 7.1.4　Sierpinski 垫

先将一个等边三角形四等分，得到四个小的全等等边三角形，挖去中间的一个，留下其余的三个小三角形. 然后对留下的三个小的三角形重复上述步骤至无穷次，就得到了 Sierpinski 垫. 每挖去一次，留下的图形的面积为原来图形面积的 3/4，周长为原来周长的 3/2，因此 Sierpinski 垫的面积为零，但其周长为无穷.

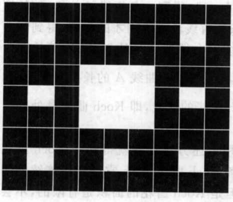


图 7.1.5　Sierpinski 地毯

如果将一个正方形九等分，挖去中间的一个小正方形. 然后对余下的八个小正方形重复上述步骤至无穷次，就得到了 Sierpinski 地毯(图 7.1.5). 同样 Sierpinski 地毯也是面积为零，周长为无穷.

四、经典分形图形的特点

从上面所介绍的几个经典分形图形，我们可以看出它们有着与我们习惯的 Euclid 空间 $\mathbf{E}^n$ 中的图形不一致的地方：

① 它们具有精细的结构，在任意小的一部分图形上也有复杂的结构；

② 它们非常不规则，以至于无法用传统的几何语言来刻画；

③ 它们的产生可以以简单的递归过程产生；

④ 它们具有某种自相似性，即将图形某一部分放大又可以得到整个图形，将整个图形缩小就得到图形的一部分.

第二节 现实中的分形

第一节所介绍的分形图形是由数学家们创造的，但在现实生活中我们也可以看见非常多的分形图形，例如树枝，河流，群山，云朵，人的血管，肺支气管等等都近似地具有分形图形所具有的特征. 即我们可以利用一些分形图形来近似描述这些物体，图 7.1.6、7.1.7 可以用来描述一棵树的树枝形状，一片树叶的形状.



图 7.1.6 分形树枝

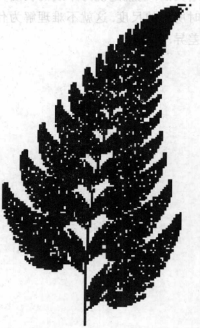


图 7.1.7 分形树叶

现实生活中一个非常著名的分形的例子是海岸线的长度.

在起伏不平的区域中选取一段海岸线来测量其长度. 显然，这个长度大于这段海岸线两个端点之间的直线距离，除非海岸线是直的，但实际情况是海岸线一般来说都是非常曲折的. 现在先采用近似测量方法. 用一个直径为 d 的圆规从海岸线的一个端点开始画圆，圆规下一步的起点恰好是上一步的终点（海岸线与圆的另一个交点），通过这样的一簇圆覆盖海岸线. 设这一簇圆的个数为 N，则海岸线的近似长度可以认为是 $L(d) = Nd$. 在近似测量中我们实际上是采用圆的直径近似代替了这个圆所覆盖的一小段曲折的海岸线. 一般而言，当减小圆的

直径 d 时，近似长度 L 会增大. Mandelbrot 在他早期最著名的一篇论文“英国的海岸线有多长?”中就研究了这一现象，得出了惊人的结论:任何海岸线都是无限长的. 这样英国的海岸线、美国海岸线和中国的海岸线具有相同的长度. 中国所有的海岸线和中国台湾的海岸线具有相同的长度.

这一现象与我们的直觉是矛盾的. 原因在于长度依赖于测量单位，以 1 千米为单位测量海岸线，得到的近似长度将短于 1 千米的迂回曲折都忽略掉了，若以 1 米为单位测量，则能测出某些被忽略掉的迂回曲折，长度将变大，测量单位进一步变小，测得的长度将愈来愈大，这些愈来愈大的长度如果趋近于一个确定值，这个极限值就是海岸线的长度. 但实际上，当测量单位变小时，所得的长度是无限增大的. 也就是说海岸线的长度，即使是近似长度，也是不确定的，它取决于测量时所用的尺度. 这就不难理解为什么不同的国家其公共边界的长度会有较大的差异.

第二章　分形的分维

维数是度量一个几何对象的一个重要的特征量. 我们习惯认为点是零维的，线是一维的，面是二维的，体是三维的. 这些维数都是整数，但对分形体而言，其维数不一定是整数. 分形几何学的主要研究工具是它的许多形式的分维.

第一节　拓　扑　维

定义　**一个几何对象的拓扑维等于确定其一个点的位置所需要的独立坐标数目，拓扑维(Topology Dimension)又叫做经验维，记为 D_T.**

在数轴上，确定一个点只需要一个坐标，因此称数轴所对应的直线是一维的. 更一般的，对二维平面上的一条曲线，记其方程为 $y = f(x)$. 要确定该曲线上的一点(x, y)的位置，需要确定 x 和 y 的值，但实际上只需要取定 x 的值，然后由 $y = f(x)$ 就得到 y 的值. 即在要确定该曲线上的一点(x, y)的位置时只需要一个独立的坐标，所以平面上曲线的拓扑维数为 1. 同理，对空间中的曲线 $x = x(t)$，$y = y(t)$，$z = z(t)$，其拓扑维也是 1. 类似的可以得到面是 2 维，体是 3 维的结论.

拓扑维是拓扑变化下的不变量. 对一个几何对象，无论通过怎样的拉伸、压缩、扭曲变成其他形状的几何对象，几何对象的拓扑维在几何对象变化前和变化后都是一样的.

我们讨论拓扑维的测量方法. 设有一边长为单位长度的正方形. 如果我们用 $r = 1/2$ 的尺子去测量，就会发现它由四个边长为 1/2 的小正方形组成，用 $N = N(r)$ 表示小正方形数目，N 和尺子长度 r 之间的关系就有 $N\left(\frac{1}{2}\right) = 4 = \frac{1}{(1/2)^2}$. 如果改用 $r = 1/10$ 的尺子来测量，则覆盖这个正方形的小正方形的个数为 $N\left(\frac{1}{10}\right) = 100 = \frac{1}{(1/10)^2}$. 虽然尺子的长度 r 不同，但是小正方形的数目 N 和 r 的负 2 次指数关系不变，即 $N(r) = \frac{1}{r^2}$. 而对于正方体，尺子的长度 r 和填充这个正方体的小正方体的数目 N 之间的关系为 $N(r) = \frac{1}{r^3}$. 一般而言，对一个维数为 D 的几何对象，尺子的长度 r 和覆盖(或填充)它的小盒子数目 N 之间

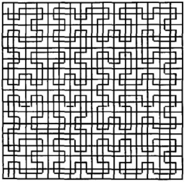


图 7.2.1 Peano 曲线

总有 $N(r)=\frac{1}{r^D}$ 成立，两边同时取对数得到 $D=-\frac{\ln N(r)}{\ln r}$. 这就是拓扑维的计算公式.

1890 年意大利数学家 Peano（皮亚诺）给出了著名的 Peano 曲线（图 7.2.1）. 可以证明 Peano 曲线经过平面上的每一个点，从而 Peano 曲线填满整个平面，故其拓扑维应该与平面的拓扑维一致，即都是 2. 但 Peano 曲线作为一条曲线，它的维数又应该是 1，这产生了矛盾. 可见拓扑维对几何对象的刻画是非常粗糙的，对于复杂的几何对象它无法进行刻画.

更直观的例子是毛线团. 从远处观察时，它是一个点，是一维的. 如果近一点它是一个球体，是三维的. 但它又是由线构成的，因此又应该是一维的. 由此，我们需要找到几何体新的维数的定义.

第二节 相 似 维

相似维是最容易理解的分形维.

根据相似性，现在再来看线段、正方形和立方体的维数. 首先，把线段、正方形和立方体的边分成两等份. 这样，线段成为一半长度的两个线段，正方形则是边长为原来边长的 1/2 的四个正方形，立方体则是边长为原来边长 1/2 的八个小立方体. 即原来的线段、正方形和立方体可以看成为分别由 2、4、8 个把全体分成 1/2 的小相似形组成. 而 2、4、8 可以记成 2^1、2^2、2^3，这里的指数 1、2、3 分别与线段、正方形和立方体的拓扑维 D_T 相一致. 一般的，如果某图形是由把全体缩小为 $1/a$ 的 a^D 个相似图形构成的，那么此指数 D 就具有维数的意义. 此维数就是相似维.

定义 如果一个几何体是由全体缩小 $1/a$ 的 b 个相似形所构成，则

$$\frac{\ln b}{\ln a}$$

叫做该几何体的相似维(similarity dimension)**，记为 D_S. a 叫做相似比例系数.**

从相似维的定义和拓扑维的计算公式可以看出，相似维是拓扑维的推广，它们的思路是一样的.

我们可以将 Cantor 集看成把全体缩小成 1/3 的两个相似形构成，所以 Cantor 集的相似维是 $\ln 2/\ln 3\approx 0.6309$. 这是一个非整数值的维数，它定量地表

示了 Cantor 集的复杂程度. 一般而言,分形体的复杂程度可以用其维数来定量地刻画. 同样的讨论可以知道,Koch 曲线的相似维是 $\ln 4/\ln 3 \approx 1.2618$, Sierpinski 垫的相似维是 $\ln 3/\ln 2 \approx 1.5850$. 特别地,Peano 曲线可以看成把全体缩小 1/2 的四个相似形构成,所以 Peano 曲线的相似维是 $\ln 4/\ln 2 = 2$. 这是一个整数值的相似维,可以看成是非整数值相似维的特殊情形.

提出相似维是把几何体的维数从整数值扩大到非整数值的划时代的进步. 但是应该看到,依据相似维的定义,只有具有严格的自相似性的分形体才具有相似维,因此它的应用范围是非常有限的.

第三节 容 量 维

容量维由于容易计算和实验测量,是一种普遍使用的维数.

定义 设几何体 F 在 n 维欧氏空间 $\mathbf{E}^n$ 中有界,$N(r)$表示覆盖 F 的边长为 r 的小立方体的最小数目,如果极限

$$\lim_{r\to 0}\frac{\ln N(r)}{\ln(1/r)} = -\lim_{r\to 0}\frac{\ln N(r)}{\ln r}$$

存在,则这个极限叫做几何体 F 的容量维(capacity dimension),**或盒维,记为 D_C.**

容量维有许多等价的定义,实际上是集合 $N(r)$的不同取法,通常有以下几种:

(1) 覆盖 F 的半径为 r 的最小闭球体数目;

(2) r 网中与 F 相交的网格数目;

(3) 覆盖 F 的直径最多为 r 的最少集合数目;

(4) 球心在 F 中,半径为 r 的相互不相交的球的最大数目.

这五种概念的 $N(r)$可以参见图 7.2.2,其中(a)与定义对应,(b)—(e)与等价定义中(1)—(4)分别对应.

对具有严格的自相似性的分形体,可以取 $r=(1/a)^k$,则 $N(r)=b^k$,故相似维和容量维相等. 因此 Cantor 集的容量维是 $\ln 2/\ln 3 \approx 0.6309$, Koch 曲线的容量维是 $\ln 4/\ln 3 \approx 1.2618$, Sierpinski 垫的容量维是 $\ln 3/\ln 2 \approx 1.5850$, Peano 曲线容量维是 $\ln 4/\ln 2 = 2$.

应该看到容量维的应用范围要比相似维的应用范围大得多. 相似维要求分形体在各个生成阶段(不同的级)具有同样的相似比例系数,但容量维对此并无要求,可以允许生产分形体时在不同的级选取不同的相似比例系数.

在实际计算容量维时,可以首先选定一个覆盖几何体 F 的立方体. 然后,任意给定一个正数 r,以 r 为边长将覆盖 F 的立方体分割成若干小立方体. 通过对

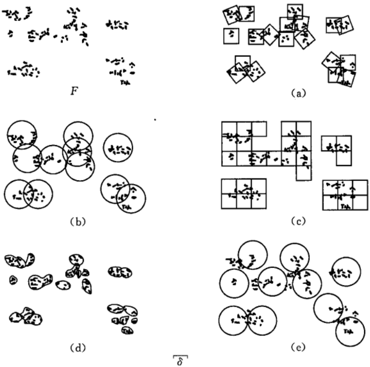


图 7.2.2　$N(r)$的不同取法

立方体内所有的点进行扫描，记录含有几何体 F 中点的小立方体数目，记为 $N(r)$，那么比值 $\ln N(r)/\ln(1/r)$ 可以看成 D_C 的近似值. 当加细分割，例如以 $r/2$ 代替 r 时，重复计算过程，得出新的比值. 如两次比值接近，则可以认为得到了几何体 F 容量维 D_C 的较精确的近似值.

第四节　信　息　维

在容量维的定义中，为了求几何体 F 的 D_C，需要知道包含 F 中点的小立方体的数目，但这样的小立方体包含多少个 F 中的点并未加考虑. 也就是说，容量维 D_C 只包含了几何体 F 的几何尺度的信息，而没有反映 F 在空间中分布的疏密信息. 为了能反映几何体 F 的空间分布信息，定义信息维.

定义 1　设几何体 F 在 n 维欧氏空间 $\mathbf{E}^n$ 中有界，$N(r)$表示覆盖 F 的边长为 r 的小立方体的最小数目. 记 p_i 为几何体 F 中的点落在第 i 个小立方体中的概率 ($i = 1, 2, \cdots, N(r)$)，令

$$I(r) = \sum_{i=1}^{N(r)} p_i \ln(1/p_i),$$

如果极限

$$\lim_{r\to 0}\frac{I(r)}{\ln(1/r)}=-\lim_{r\to 0}\frac{I(r)}{\ln r}$$

存在，则这个极限叫做几何体 F 的信息维(information dimension)**，记为 D_I.**

如果所有的小立方体以相等的概率包含几何体 F 中的点，则 $p_i=1/N(r)$，于是 $I(r)=\ln N(r)$. 这样信息维 D_I 便与容量维 D_C 一致. 可见，容量维 D_C 是信息维 D_I 的一种特殊情形，或信息维 D_I 是容量维 D_C 的推广. 计算概率 p_i 的简单方法是用落在第 i 个小立方体中的点的频率代替概率，因此在计算时要求点的数目比较大.

还可以将信息维 D_I 推广到高次信息维 D_q.

定义 2　设几何体 F 在 n 维欧氏空间 $\mathbf{E}^n$ 中有界，$N(r)$ 表示覆盖 F 的边长为 r 的小立方体的最小数目. 记 p_i 为几何体 F 中的点落在第 i 个小立方体中的概率 $(i=1,2,\cdots,N(r))$，令

$$I_q(r)=\frac{1}{1-q}\ln\sum_{i=1}^{N(r)}p_i^q,$$

如果极限

$$\lim_{r\to 0}\frac{I_q(r)}{\ln r}$$

存在，则这个极限叫做几何体 F 的 q 次信息维，记为 D_q.

高次信息维又叫广义维，它是将信息维广义化后的量. 如果考虑让 $q\to 1$，因为 $p_i^q=p_i\times p_i^{q-1}=p_i\exp[(q-1)\ln p_i]$ 及 $[(q-1)\ln p_i]$ 的值很小，可以近似的将 p_i^q 表示为 $p_i[1+(q-1)\ln p_i]$，再应用 L'Hospital 法则得

$$\begin{aligned}\lim_{q\to 1}I_q(r)&=\lim_{q\to 1}\frac{1}{1-q}\ln\sum_{i=1}^{N(r)}p_i^q\\&=\lim_{q\to 1}\frac{\ln\sum_{i=1}^{N(r)}[p_i+(q-1)p_i\ln p_i]}{1-q}\\&=-\sum_{i=1}^{N(r)}p_i\ln p_i=I(r).\end{aligned}$$

此时得到的广义维就是信息维，即 $D_q|_{q=1}=D_I$.

如果让 $q\to 0^+$，则此时的广义维就是容量维，即 $D_q|_{q=0}=D_C$. 这是因为

$$\lim_{q\to 0^+}p_i^q=\begin{cases}0, & 当\ p_i=0;\\1, & 当\ p_i\neq 0,\end{cases}$$

即 $\sum_{i=1}^{N(r)} p_i^0$ 只能等于含有几何体 F 中点的小立方体数目.

高次信息维 D_q 有一个重要的性质，q 越大，D_q 越小，即通常有下面的不等式

$$D_{q_1} \geqslant D_{q_2} \qquad (\text{当 } q_1 \leqslant q_2 \text{ 时}).$$

根据这个不等式，取 $q = 0$ 和 1，则有 $D_C \geqslant D_I$.

第五节　关　联　维

相似维、容量维和信息维对低维空间的几何对象非常有用，但是在高维空间中的应用却受到很大的限制，因此从理论上讲迫切需要定义其他的分维数. 自然界的事物十分复杂，目前我们完全认识的只是其中较简单的那一部分，对复杂系统的组成、状态等还缺乏透彻的了解. 但可以肯定的是，复杂系统的状态空间、参数空间等的维数都是高维的，其维数可能成百上千. 对于如此复杂的系统，能否寻找一种方法降低它的维数呢？甚至用一种简单的时间序列来映射系统的运动变化特征呢？这就产生了关联维. 关联维常常用来刻画时间序列 $x_1, x_2, \cdots, x_n$ 的自相似性.

首先由时间序列 $x_1, x_2, \cdots, x_n$ 构造出 m 维相空间(m 称为**嵌入维**). 常用的方法是时间差法，即按一定的时间间隔 p 从时间序列中取出数来作为分量构造向量：

$$y_i = (x_i, x_{i+p}, x_{i+2p}, \cdots, x_{i+(m-1)p}) \ (i = 1, 2, \cdots).$$

构造好向量 y_i 后，定义向量 y_i 和 y_j 之间的距离：

$$r_{ij} = \max_{1 \leqslant k \leqslant m} | y_{ik} - y_{jk} |,$$

其中，y_{ik} 表示向量 y_i 的第 k 个分量，即 $y_{ik} = x_{i+(k-1)p}$.

如果 y_i 和 y_j 之间的距离小于给定的标度 r 就称这两个向量是有关联的. 若总共构造了 N 个向量，则有 N^2 对向量可能相关联. 对于给定的标度 r，统计相空间中相关联的向量对数目所占的比例：

$$C(r) = \frac{1}{N^2} \sum_{i,j=1}^{N} H(r - r_{ij}),$$

其中，函数 $H(x)$ 为阶跃函数，即

$$H(x) = \begin{cases} 1, & \text{当 } x > 0; \\ 0, & \text{当 } x \leqslant 0. \end{cases}$$

如果标度 r 取得太大，任何一对向量都是关联的，$C(r)=1$，取对数后为 0；如果标度 r 取得合适的话，就可以定义关联维了.

定义　如果极限

$$\lim_{r\to 0}\frac{\ln C(r)}{\ln r}$$

存在，则这个极限叫做时间序列 $x_1, x_2, \cdots, x_n$ **的关联维**(correlation dimension)，**记为** D_{Co}. **其中，**$C(r)$**的定义如前所述.**

如果标度 r 取得太小，低于环境噪音和测量误差造成的向量差异，则上式计算得到的就不是关联维 D_{Co}，而是嵌入维 m 了. 因此在实际计算中，往往采用下面的计算步骤：

(1) 给定一个时间间隔 p，给定一个嵌入维 m，给出并变化标度 r，计算出点列$\{(\ln r, \ln C(r))\}$；

(2) 作出点列$\{(\ln r, \ln C(r)\}$的点图；

(3) 在 $r=0$ 附近求点列$(\ln r, \ln C(r))$的拟合直线，得到拟合直线的斜率；

(4) 变化嵌入维 m，重复上面的过程，直到拟合直线的斜率不再发生变化.

此时的嵌入维 m 称为饱和嵌入维 m_c；此时的拟合直线斜率就是关联维 D_{Co}.

饱和嵌入维 m_c 是一个重要的参数. 如果饱和嵌入维 m_c 存在，就说明该时间序列具有分形性质，否则该时间序列是随机序列，不具有分形性质.

第六节　Hausdorff 维

波恩大学数学家 Hausdorff 1919 年从测度的角度引入了 Hausdorff 维. 在讨论 Hausdorff 维之前，先来看一下 Hausdorff 测度.

定义 1　设 F **为** n **维欧氏空间** $\mathbf{E}^n$ **的几何体，**$d(x, y)$**为两点间的距离函数，则**

$$\sup\{d(x, y)\mid x, y\in F\}$$

叫做几何体 F 的直径，记为$|F|$.

定义 2　设 F **为** n **维欧氏空间** $\mathbf{E}^n$ **的几何体，设有可数个直径不超过** $\delta\,(\delta>0)$ **的几何体** F_i**，即** $0<|F_i|\leqslant\delta$**，使得** $F\subseteq\bigcup_{i=1}^{\infty}F_i$ **，则说**$\{F_i\}$**为** F **的一个** δ **覆盖.**

一般而言，F 的 δ 覆盖是不止一个的.

定义 3　设 s **是一个非负数，定义**

$$H_\delta^s=\inf\Big\{\sum_{i=1}^{\infty}|F_i|^s \;\Big|\; \{F_i\}\ \text{为}\ F\ \text{的}\ \delta\ \text{覆盖}\Big\}.$$

如果极限 $\lim_{\delta\to 0^+} H_\delta^s(F)$ **存在，则把这个极限叫做** F **的** s **维 Hausdorff 测度，记** $H^s(F)=\lim_{\delta\to 0^+} H_\delta^s(F)$.

实际上，$\lim_{\delta\to 0^+} H_\delta^s(F)$总是存在的，但极限值可以是(并且通常是)0 或$\infty$.

如果 $F=\varnothing$，则 $H^s(F)=0$；如果 $E\subseteq F$，则 $H^s(E)\leqslant H^s(F)$. 可以证明，如果$\{F_i\}$为任何可数的分离 Borel 集序列，则 $H^s(\bigcup_{i=1}^{\infty} F_i)=\sum_{i=1}^{\infty} H^s(F_i)$. 测度基本上只是赋予集合(这里指几何体 F)以数值"大小"的一种方式，如果集合是以合理的方式分解为有限或可数个部分，则整体的数值应该是各部分数值之和. 这就是我们为什么把 $H^s(F)$叫做测度. Hausdorff 测度推广了长度、面积和体积等类似的概念.

Hausdorff 测度具有如下性质：

(1) 设 $F\subset \mathbf{E}^n$，$\lambda>0$，则有定比性质：

$$H^s(\lambda F)=\lambda^s H^s(F),$$

其中，$\lambda F=\{\lambda x \mid x\in F\}$.

(2) 设 $F\subset \mathbf{E}^n$，$f:F\to \mathbf{E}^m$ 为一映射，且对正值常数 c 和α，有 Holder 条件：

$$|f(x)-f(y)|\leqslant c\,|x-y|^\alpha,\ \forall\, x,\ y\in F,$$

则对任意 s 有

$$H^{s/\alpha}\varphi(f(F))=c^{s/\alpha}H^s(F).$$

所以，如果 $|f(x)-f(y)|=|x-y|$，则有 $H^s(f(F))=H^s(F)$，即 Hausdorff 测度对平移和旋转保持不变.

下面来讨论 Hausdorff 维. 从 $H_\delta^s(F)$的定义可以看出，对任意给定的几何体 F 和$\delta<1$，$H_\delta^s(F)$对 s 是不增的，因此 $H^s(F)$对 s 也是不增的. 如果令 $t>s$，$\{F_i\}$是 F 的一个 δ 覆盖，则有

$$\sum_i |F_i|^t\leqslant \delta^{t-s}\sum_i |F_i|^s,$$

取下确界，得 $H_\delta^t(F)\leqslant \delta^{t-s}H_\delta^s(F)$. 让 $\delta\to 0^+$，对于 $t>s$，若 $H^s(F)<\infty$，则 $H^t(F)=0$. 所以 $H^s(F)$关于 s 的图应该如图 7.2.3 所示. 这表明存在 s 的一个临界值，使得 $H^s(F)$从∞"跃"到 0. 这个临界值就叫 F 的 Hausdorff 维.

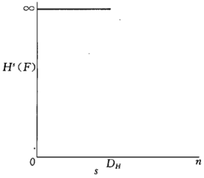


图 7.2.3　$H^s(F)$对 s 的图

定义 4　设 F **为** n **维欧氏空间** $\mathbf{E}^n$ **中的**

几何体，

$$\inf\{s \mid H^s(F) = 0\} = \sup\{s \mid H^s(F) = \infty\}$$

叫做 F 的Hausdorff 维，记为 D_H.

从 Hausdorff 维 D_H 的定义可以知道，

$$H^s(F) = \begin{cases} \infty, & \text{当 } s < D_H \text{ 时；} \\ 0, & \text{当 } s > D_H \text{ 时.} \end{cases}$$

如果 $s = D_H$，则 $H^s(F)$可以为 0，或∞，或一个有限的正数.

从 Hausdorff 测度的性质可以得到 Hausdorff 维 D_H 的许多性质：

(1) 如果 $F \subset \mathbf{E}^n$ 为开集，则 $D_H(F) = n$，这是因为 F 中包含一个 n 维体积的球体.

(2) 如果 F 为 $\mathbf{E}^n$ 中的 m 维光滑曲面，则 $D_H(F) = m$.

(3) 如果 $E \subset F$，则 $D_H(E) < D_H(F)$.

(4) 如果$\{F_i\}$为可数的序列，则 $D_H(\bigcup_{i=1}^{\infty} F_i) = \sup_{1 \leqslant i < \infty} \{D_H(F_i)\}$.

(5) 如果 F 是可数的，则 $D_H(F) = 0$.

(6) 如果 $F \subset \mathbf{E}^n$，映射 $f: F \to \mathbf{E}^m$ 满足 Holder 条件：

$$| f(x) - f(y) | \leqslant c \mid x - y \mid^{\alpha}, \qquad \forall x, y \in F,$$

则 $D_H(f(F)) \leqslant \frac{1}{\alpha} D_H(F)$.

(7) 如果 $F \subset \mathbf{E}^n$，映射 $f: F \to \mathbf{E}^m$ 满足双 Lipschitz 条件：

$$c_1 \mid x - y \mid \leqslant \mid f(x) - f(y) \mid \leqslant c_2 \mid x - y \mid, \quad \forall x, y \in F, 0 < c_1 < c_2 < \infty,$$

则 $D_H(f(F)) = D_H(F)$.

在计算中，常常采用基于"用尺度进行度量"的设想，并且忽略小于尺度时的不规则性，然后查看当尺度趋向于 0 时这些测量值的变化. 这实际上就是前面介绍过的容量维，所以容量维应该是 Hausdorff 维的变形，因此在有些文献中直接就将容量维叫做 Hausdorff 维. 容量维的计算是非常方便的，不需要用到测度理论，但也正因为 Hausdorff 维是用测度来定义的，所以在严谨的理论研究上就非常重要了.

第七节　分形的概念

目前关于分形的定义还没有取得一致意见. Mandelbrot 1982 年曾经给出分形的一个原始定义，把 Hausdorff 维 D_H 严格大于拓扑维 D_T 的集合叫做分形.

但由于发现该定义不能包含已经发现的部分分形图形，四年后他又提出了一个实用的定义，把组成部分以某种方式与整体相似的几何体叫做分形. 这个定义很通俗且直观，很受大家的欢迎. 事实上，人们一直期望能够像通常的数学名词一样给出分形一个严格的定义，但是现在还没有这样的分形的定义。不过这并不影响对分形的研究和应用，就像现在还没有生命的定义，但是对生命科学的研究一直在不断地深入。

这里的“某种方式”是指“自相似”，即在任何尺度下观察某一几何体时都可以得到相同的图形，或者称其为“无特征尺度”. 所谓特征尺度是指某一种事物在空间或时间尺度上具有的特定的数量级. 例如，描述两个恒星体的距离的特征尺度用光年，用千米是不可以的；描述两个城市的距离的特征尺度一般用千米，用光年或厘米就没有多少意义了；而要描述一本书的大小用厘米或毫米才是合适的. 但是对于有些几何体，它本身包含很多层次，在不同的层次上观察其形状却是彼此相似的，我们也说这些几何体无特征尺度，即在一定范围内用任何尺度考察都有相同或相似的结果. 对于现实中的海岸线在一定范围内是无尺度特征的. 当在一张比例尺为百万分之一的地图上看一个海湾时，如果将地图的比例尺放大为十万分之一，就会看见在大海湾的背景下还有一些小的海湾，继续放大比例尺则在小海湾的背景下还有更小的海湾出现……这就告诉我们在一定的范围内，无论用什么样的比例尺，都可以看见形状相似的海湾，即海岸线是一种无特征尺度的自相似分形体.

分形的特征除了相似性外，还有分维性，即其维数一般是非整数. 分数维的产生是由几何体的非规则性、内部非均匀性与破碎性决定的. 对于一几何体，其整体越不规则，内部越不均匀或越破碎，则分维值越大. 分维值的大小，反映了几何体整体的不规则程度、内部的不均匀程度或破碎程度. 需要说明的是，分维还有许多其他的定义，在不同的领域中，根据研究对象不同的特点及不同的研究目的往往会采用不同的分维定义.

第三章　分形的生成

分形具有精细的结构，在任意小的一部分上也有复杂的结构，而且它们非常不规则，以至于无法用传统的几何语言来刻画. 但是分形集合应该而且必须能够刻画出来. 许多分形可以看成是由一些与整体以某种方式相似的部分组成，因此可以采用简单的递归过程来产生. 1981 年 Hutchinson 首先提出可以用压缩映射的不动点理论来产生分形，继而 1985 年 M. Barmsley 发明了迭代函数系(iterated function system，简记为 IFS). IFS 是描绘分形的强有力的数学方法.

第一节　分形空间

分形理论与应用问题的研究总是在一个假定的理想空间中进行的. 一个完备的度量空间常常能够满足我们的要求. 这个完备的度量空间就是我们要讨论的分形空间.

设(X, d)是一个完备的度量空间，记 $F(X)$为 X 的全体非空紧子集组成的空间，即 X 的一个非空紧子集是 $F(X)$中的一个点. 下面定义 $F(X)$中的度量.

定义 1　**设(X, d)是一个完备的度量空间，$x \in X$，集合 $B \in F(X)$，**

$$d(x, B) = \min\{d(x, y) \mid y \in B\}$$

叫做点 x 到集合 B 的距离，记为 $d(x, B)$.

定义 2　**设(X, d)是一个完备的度量空间，集合 A、$B \in F(X)$，称**

$$d(A, B) = \max\{d(x, B) \mid x \in A\},$$

为集合 A、B 的距离，记为 $d(A, B)$.

显然，$d(A, A) = \max\{d(x, A) \mid x \in A\} = 0$.

当 $d(A, B) = 0$ 时，对 A 中任意一点 x，有 $d(x, B) = 0$，即存在 $y \in B$，使得 $d(x, y) = 0$，所以 $y = x$，从而 $x \in B$. 这就证明了如果 $d(A, B) = 0$，则 $A \subset B$.

如果集合 $A, B \in F(X)$，一般来说 $d(A, B) \neq d(B, A)$，因此 $d(A, B)$不能作为 $F(X)$中的度量.

如果集合 $A, B, C \in F(X)$，则 $d(A, B) \leqslant d(A, C) + d(C, B)$. 这是因为

$$d(x, B) = \min\{d(x, y) \mid y \in B\}$$
$$\leqslant \min\{d(x, z) + d(z, y) \mid y \in B\}, \ \forall z \in C,$$
$$= d(x, z) + \min\{d(z, y) \mid y \in B\}.$$

令 $d(x, z_0) = \min\{d(x, z) \mid z \in C\}$，则

$$d(x, B) \leqslant d(x, z_0) + \min\{d(z_0, y) \mid y \in B\}$$
$$= d(x, C) + \min\{d(z_0, y) \mid y \in B\}$$
$$\leqslant d(x, C) + \max\{\min\{d(z, y) \mid y \in B\} \mid z \in C\}$$
$$= d(x, C) + d(C, B),$$

两边对 $x \in X$ 取最大值，得 $d(A, B) \leqslant d(A, C) + d(C, B)$.

定义 3 设(X, d)是一个完备的度量空间，集合 A、$B \in F(X)$，

$$h(A, B) = d(A, B) \vee d(B, A),$$

叫做集合 A，B 的Hausdorff 距离.

定理 1 h **是空间$F(X)$中的一个度量.**

证 令集合 $A, B, C \in F(X)$.

(1) 显然，$h(A, B) \geqslant 0$, $h(A, A) = d(A, A) \vee d(A, A) = d(A, A) = 0$;

$h(A, B) = 0$ 时，$d(A, B) = 0$ 且 $d(B, A) = 0$. 由 $d(A, B) = 0$ 得 $A \subset B$，由 $d(B, A) = 0$ 得 $B \subset A$，故 $A = B$.

(2) 由定义，$h(A, B) = h(B, A)$.

(3) 由 $d(A, B) \leqslant d(A, C) + d(C, B)$ 及 $d(B, A) \leqslant d(B, C) + d(C, A)$ 得

$$h(A, B) = d(A, B) \vee d(B, A)$$
$$\leqslant [d(A, C) + d(C, B)] \vee [d(B, C) + d(C, A)]$$
$$= d(A, C) \vee d(C, A) + d(C, B) \vee d(B, C)$$
$$= h(A, C) + h(C, B).$$

今后度量空间$(F(X), h)$叫做分形空间，可以证明$(F(X), h)$为完备空间.

第二节 压缩映射与 IFS

设 f 是度量空间 X 上的一个变换，记 $y = f(x)$. 对度量空间 X 上的某一点 x 代入得到 y 后，以 y 为新的 x 代入再得到新的 y 值，不断的重复这一操作，记 $f^0(x) = x$, $f^1(x) = f(x)$, $f^2(x) = f \cdot f(x) = f(f(x))$, $\cdots$, $f^{n+1}(x) = f \cdot f^n(x) = f(f^n(x))(n = 0, 1, 2, \cdots)$. 变换 f^n 叫做度量空间 X 上的一个迭代.

定义 1 设变换 $W:\mathbf{E}^2\to\mathbf{E}^2$ 具有形式

$$W\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}e\\f\end{bmatrix},$$

其中(x, y)为 $\mathbf{E}^2$ 中的点，a, b, c, d, e, f 均为实数，则 W 叫做一个(二维)仿射变换.

当记 $\mathbf{E}^2$ 中的点为 x 时，上式常记为

$$W(x)=Ax+t,$$

其中 $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, $t=\begin{bmatrix}e\\f\end{bmatrix}$.

有四种平面仿射变换具有明显的几何意义，记

$$A_s=\begin{bmatrix}s&0\\0&s\end{bmatrix},\ A_t=\begin{bmatrix}1&0\\0&t\end{bmatrix},\ A_u=\begin{bmatrix}1&u\\0&1\end{bmatrix},\ A_\theta=\begin{bmatrix}\cos\theta&\sin\theta\\\sin\theta&\cos\theta\end{bmatrix},$$

A_s 为缩放变换，A_t 为伸长变换，A_u 为剪切变换，A_θ 为旋转变换(图 7.3.1).

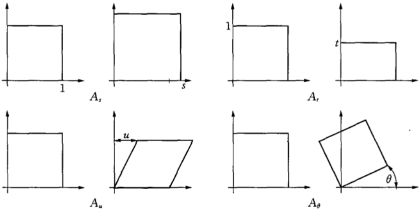


图 7.3.1 四种平面仿射变换

定义 2 设 f 为度量空间(X, d)上的变换，如果存在常数 $0\leqslant s<1$，使得 f 满足

$$d(f(x), f(y))\leqslant sd(x, y),\ \forall x, y\in X,$$

则 f 叫做压缩映射，常数 s 叫做压缩因子.

IFS 是由一个完备的度量空间 X, X 上有限个压缩映射 $W_n(n=1, 2, \cdots, N)$ 及分别对应的压缩因子 s_n 所构成的，用$\{X; W_n, n=1, 2, \cdots, N\}$ 表示 IFS，其压缩因子为 $s=\max\{s_n \mid n=1, 2, \cdots, N\}$.

定理 1　设$\{X; W_n, n=1, 2, \cdots, N\}$是以$s$为压缩因子的 IFS,定义分形空间$(F(X), h)$上的变换$W$为

$$W(B)=\bigcup_{n=1}^{N} W_n(B), \ \forall B \in F(X),$$

则W为完备空间$(F(X), h)$上以s为压缩因子的压缩映射,它唯一的不动点P满足

$$P=W(P)=\bigcup_{n=1}^{N} W_n(P),$$

且$P=\lim\limits_{n \to \infty} W^n(B), \ \forall B \in F(X)$.

定理 1 的证明略.

定理 1 中的不动点P称为 IFS 的吸引子,是一个分形体,叫做自仿射分形. IFS 的基本思想并不复杂,它认定几何对象的全貌与局部,在仿射变换的意义下,具有自相似结构. 这样一来,几何对象的整体被定义之后,选定若干仿射变换,将整体形态变换到局部. 并且这一过程可以不断地进行下去,直到得到满意的结果.

例 1　设d为欧氏空间$\mathbf{E}^1$上的距离. 考虑 IFS$\{\mathbf{R}; W_1, W_2\}$,其中$W_1(x)=\frac{1}{3}x$, $W_2(x)=\frac{1}{3}x+\frac{2}{3}$, $B_0=[0, 1]$,问$P=\lim\limits_{n \to \infty} W^n(B_0)$为何物? 压缩因子$s$的值为多少?

解　由"压缩映射是连续的"知道对$W=W_1 \cup W_2$是连续的,因此只要每次映射两端点就行了.

$$W_1(0)=0, \ W_1(1)=\frac{1}{3} \Rightarrow W_1(B_0)=\left[0, \frac{1}{3}\right];$$

$$W_2(0)=\frac{2}{3}, \ W_2(1)=1 \Rightarrow W_2(B_0)=\left[\frac{2}{3}, 1\right];$$

所以$B_1=W(B_0)=W_1(B_0) \cup W_2(B_0)=\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$.

$$W_1(0)=0, \ W_1\left(\frac{1}{3}\right)=\frac{1}{9} \Rightarrow W_1\left(\left[0, \frac{1}{3}\right]\right)=\left[0, \frac{1}{9}\right];$$

$$W_2(0)=\frac{2}{3}, \ W_2\left(\frac{1}{3}\right)=\frac{7}{9} \Rightarrow W_2\left(\left[0, \frac{1}{3}\right]\right)=\left[\frac{6}{9}, \frac{7}{9}\right];$$

$$W_1\left(\frac{2}{3}\right)=\frac{2}{9}, \ W_1(1)=\frac{1}{3} \Rightarrow W_1\left(\left[\frac{2}{3}, 1\right]\right)=\left[\frac{2}{9}, \frac{3}{9}\right];$$

$$W_2\left(\frac{2}{3}\right)=\frac{8}{9}, \ W_2(1)=1 \Rightarrow W_2\left(\left[\frac{2}{3}, 1\right]\right)=\left[\frac{8}{9}, 1\right];$$

所以 $W_1(B_1)=\left[0,\frac{1}{9}\right]\cup\left[\frac{2}{9},\frac{3}{9}\right]$；$W_2(B_1)=\left[\frac{6}{9},\frac{7}{9}\right]\cup\left[\frac{8}{9},1\right]$；

$$B_2=W(B_1)=W_1(B_1)\cup W_2(B_1)$$
$$=\left[0,\frac{1}{9}\right]\cup\left[\frac{2}{9},\frac{3}{9}\right]\cup\left[\frac{6}{9},\frac{7}{9}\right]\cup\left[\frac{8}{9},1\right].$$

显然，$P=\lim\limits_{n\to\infty}W^n(B_0)$ 就是经典的 Cantor 集(图 7.3.2).

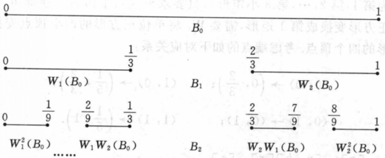


图 7.3.2　产生 Cantor 集的 IFS 示意图

又　$d(W_1(x),W_1(y))=\left|\frac{1}{3}x-\frac{1}{3}y\right|=\frac{1}{3}\mid x-y\mid=\frac{1}{3}d(x,y)$,

$d(W_2(x),W_2(y))=\left|\frac{1}{3}x+\frac{2}{3}-\frac{1}{3}y-\frac{2}{3}\right|=\frac{1}{3}\mid x-y\mid=\frac{1}{3}d(x,y)$,

即 $s_1=\frac{1}{3}$, $s_2=\frac{1}{3}$，所以 $s=\max\{s_1,s_2\}=\frac{1}{3}$.

前面的例子说明了如何由给定的 IFS 求其吸引子，反过来，如果已经给出了分形图形，如何由分形图形求出 IFS 呢？简单的方法是找出原像 A_0 和第一次迭代后的图形 A_1，分析两者之间的关系，确定出几个对应的仿射变换即可.

例 2　如图 7.3.3(a)所示，A_0 是二维 Euclid 空间的一个单位正方形，将该

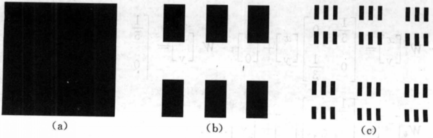


图 7.3.3　自仿射分形

正方形水平方向五等份，去掉第二、第四两部分，同时在铅直方向上将正方形三等份，去掉第二部分，留下长宽比为 3:5 的六个小矩形(图 7.3.3(b)). 然后对每个小矩形重复上面的操作(图 7.3.3(c))下去就会产生分形体. 求产生该分形图形的 IFS.

解 由于第一次迭代后有六个小的矩形，所以 IFS 应该含有六个变换将正方形分别变换成六个小的矩形. 设 IFS 为$\{\mathbf{R}^2; W_1, W_2, W_3, W_4, W_5, W_6\}$，其中 W_i 将正方形变换成第 i 小矩形(在图 7.3.2(b)中按照从左到右，从上到下分别命名为第 1，第 2，…，第 6 小矩形)，只要求出这六个仿射变换即可. 要使 W_1 将单位正方形变换成第 1 矩形，需要 W_1 将单位正方形的四个顶点变换成第 1 个小矩形的四个顶点. 考虑端点的如下对应关系：

$$(0,0)\to\left(0,\frac{2}{3}\right);\qquad (1,0)\to\left(\frac{1}{5},\frac{2}{3}\right);$$

$$(0,1)\to(0,1);\qquad (1,1)\to\left(\frac{1}{5},1\right).$$

设 $W_1\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}e\\f\end{bmatrix}$，将上面的条件代入可以解得

$$a=\frac{1}{5},\ b=0,\ c=0,\ d=\frac{1}{3},\ e=0,\ f=\frac{2}{3},$$

即

$$W_1\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}0\\\frac{2}{3}\end{bmatrix}.$$

同理可以求出其余的五个仿射变换，

$$W_2\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}\frac{2}{5}\\\frac{2}{3}\end{bmatrix};\quad W_3\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}\frac{4}{5}\\\frac{2}{3}\end{bmatrix};$$

$$W_4\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}0\\0\end{bmatrix};\quad W_5\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}\frac{2}{5}\\0\end{bmatrix};$$

$$W_6\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{5}&0\\0&\frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}\frac{4}{5}\\0\end{bmatrix}.$$

需要说明的是，这样的仿射变换不唯一，如可以用

$$W_1'\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0 & -\frac{1}{5}\\ -\frac{1}{3} & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}\frac{1}{5}\\ 1\end{bmatrix}$$

代替 W_1 同样可以将单位正方形变换为第 1 小矩形，只是此时的顶点对应关系变为

$$(0,\ 0)\to\left(\frac{1}{5},\ 1\right);\qquad (1,\ 0)\to\left(\frac{1}{5},\ \frac{2}{3}\right);$$

$$(0,\ 1)\to(0,\ 1);\qquad (1,\ 1)\to\left(0,\ \frac{2}{3}\right).$$

第三节　分形曲面插值

分形体可以看成是一个 IFS 的吸引子. 分形插值方法是依据 IFS 而建立起来的一种新的数值方法，可以用于模拟具有分形性质的几何形体. 本节只介绍分形曲面插值的迭代算法.

分形曲面插值的思想是让整个区域与每一个小区域都相似. 图 7.3.4 中，假设整个区域是一个包含四个小正方形的大正方形，将各点编号 1，2，…，9. 当进行分形插值时，大正方形各点应该与第一个小正方形的各点有对应关系：1↔1，2↔12，3↔2，4↔21，5↔22，6↔23，7↔4，8↔32，9↔5. 实际上只需要两个压缩映射 $L_1(x)=x/2$，$K_1(y)=y/2$ 就可以做到这一点. 当再构造出函数 $f_1(x,\ y,\ z)$ 来计算各点上的函数值时，函数组 $L_1(x)$，$K_1(y)$，$f_1(x,\ y,\ z)$ 就完成了对第一个小正方形的插值. 同样，在其余三个小正方形中进行插值时也要分别构造出三组相应的函数组.

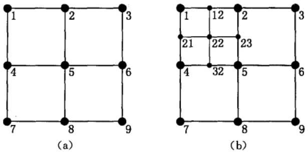


图 7.3.4　分形插值示意图

一般的,设区间 $I=[a, b]$,区间 $J=[c, d]$;设平面区域 $D=I\times J=\{(x, y)\mid a\leqslant x\leqslant b,\ c\leqslant y\leqslant d\}$,以 Δx, Δy 为步长,将 D 剖分为网格

$$a=x_0<x_1<\cdots<x_N=b,$$

$$c=y_0<y_1<\cdots<y_M=d.$$

在网格上给定一组数据 $(x_n, y_m, z_{n,m})$, $n=0, 1, 2, \cdots, N$; $m=0, 1, 2, \cdots, M$. 欲构造一个分形插值函数 $f: D\to\mathbf{R}$ 满足

$$f(x_n, y_m)=z_{n,m},\quad n=0, 1, 2, \cdots, N;\ m=0, 1, 2, \cdots, M.$$

记 $I_n=[x_{n-1}, x_n]$, $J_m=[y_{m-1}, y_m]$, $D_{n,m}=I_n\times J_m(n=0, 1, 2, \cdots, N; m=0, 1, 2, \cdots, M)$. 令 $L_n: I\to I_n$, $K_m: J\to J_m$ 为压缩变换,并满足边界条件

$$\begin{cases}L_n(x_0)=x_{n-1},\ L_n(x_N)=x_n,\\ K_m(y_0)=y_{m-1},\ K_m(y_M)=y_m,\\ |L_n(\xi_1)-L_n(\xi_2)|<l_1|\xi_1-\xi_2|,\\ |K_m(\eta_1)-K_m(\eta_2)|<l_2|\eta_1-\eta_2|,\end{cases}$$

其中, $\xi_1, \xi_2\in I$; $\eta_1, \eta_2\in J$; $0\leqslant l_1, l_2<1$.

一般采用线性函数来构造 L_n 和 K_m. 令 $L_n(x)=a_nx+b_n$, $K_m(y)=c_my+d_m$,则由边界条件可以得到

$$a_n=\frac{x_n-x_{n-1}}{x_N-x_0},\ b_n=\frac{x_{n-1}x_N-x_nx_0}{x_N-x_0},$$

$$c_m=\frac{y_m-y_{m-1}}{y_M-y_0},\ d_m=\frac{y_{m-1}y_M-y_my_0}{y_M-y_0}.$$

设计算插值点函数值的 $F_{n,m}: D_{n,m}\to[z_{\min}, z_{\max}]$ 是连续的,并满足端点条件

$$\begin{cases}F_{n,m}(x_0, y_0, z_{0,0})=z_{n-1,m-1},\\ F_{n,m}(x_N, y_0, z_{N,0})=z_{n,m-1},\\ F_{n,m}(x_0, y_M, z_{0,M})=z_{n-1,m},\\ F_{n,m}(x_N, y_M, z_{N,M})=z_{n,m}.\end{cases}$$

又 $F_{n,m}$ 是压缩的,故

$$|F_{n,m}(\xi_1, \eta_1, \tau_1)-F_{n,m}(\xi_2, \eta_2, \tau_2)|<l_3|\tau_1-\tau_2|,$$

其中,

$$\xi_1, \xi_2\in I;\ \eta_1, \eta_2\in J,\ \tau_1, \tau_2\in[z_{\min}, z_{\max}],\ 0\leqslant l_3<1.$$

假设

$$F_{n,m}(x, y, z) = e_{n,n}x + f_{n,m}y + g_{n,m}xy + \alpha_{n,m}z + k_{n,m},$$

则 $F_{n,m}(x, y, z)$中各常数可以通过边界条件来确定，只有 $\alpha_{n,m}$为自由参量，但必须满足其绝对值小于 1. 由端点条件可以得出参数 $e_{n,m}$，$f_{n,m}$，$g_{n,m}$，$k_{n,m}$的取值. 记

$$h_1 = z_{n-1,m-1} - z_{n-1,m} - z_{n,m} - \alpha_{n,m}(z_{0,0} - z_{N,0} - z_{0,M} + z_{N,M}),$$

$$h_2 = x_0y_0 - x_Ny_0 - x_0y_M + x_Ny_M,$$

$$h_3 = z_{n-1,m-1} - z_{n,m-1} - \alpha_{n,m}(z_{0,0} - z_{N,0}) - g_{n,m}(x_0y_0 - x_Ny_0),$$

$$h_4 = z_{n-1,m-1} - z_{n-1,m} - \alpha_{n,m}(z_{0,0} - z_{0,M}) - g_{n,m}(x_0y_0 - x_0y_M),$$

则有

$$g_{n,m} = h_1/h_2,$$

$$e_{n,m} = h_3/(x_0 - x_N),$$

$$f_{n,m} = h_4/(y_0 - y_M),$$

$$k_{n,m} = z_{n,m} - e_{n,m}x_N - f_{n,m}y_N - \alpha_{n,m}z_{N,M} - g_{n,m}x_Ny_M.$$

记 $S = D \times [z_{\min}, z_{\max}]$，定义 $G_{n,m}(x, y, z)$如下

$$G_{n,m}(x, y, z) = \begin{cases} \frac{1}{2}[F_{n,m}(x, y, z) + F_{n+1,m}(x_0, y, z)], \\ \qquad \text{当 } x = x_N \text{ 时}; n = 1, 2, \cdots, N-1; m = 1, 2, \cdots, M; \\ \frac{1}{2}[F_{n,m}(x, y, z) + F_{n,m+1}(x, y_0, z)], \\ \qquad \text{当 } y = y_M \text{ 时}; n = 1, 2, \cdots, N; m = 1, 2, \cdots, M-1; \\ F_{n,m}(x, y, z), \text{其他}, n = 1, 2, \cdots, N; m = 1, 2, \cdots, M. \end{cases}$$

在 S 上定义函数组 $W_{n,m}(x, y, z)$为

$$W_{n,m}(x, y, z) = \{L_n(x), K_m(y), G_{n,m}(x, y, z)\},$$

$$n = 0, 1, 2, \cdots, N; m = 0, 1, 2, \cdots, M,$$

则得到 S 上的一个 IFS：

$$\{S; W_{n,m}, n = 1, 2, \cdots, N, m = 1, 2, \cdots, M\}.$$

当自由参量 $\alpha_{n,m}$的绝对值小于 1 时，存在惟一的 $G \subset S$，使得 $G = \bigcup_{n=1}^{N} \bigcup_{m=1}^{M} W_{n,m}(G)$ 为唯一的吸引子，G 就是满足条件

$$f(x_n, y_m) = z_{n,m}, n = 0, 1, 2, \cdots, N; m = 0, 1, 2, \cdots, M$$

的分形插值曲面.

第四章　分形的应用

分形理论从其诞生之日起，就在许多学科领域得到了广泛的应用. 地学中的许多事物都十分复杂，是非线性和不规则的，广泛地存在无序、混乱、不规则和不光滑的复杂现象. 因此，作为非线性科学的一个重要分支，分形理论从诞生到发展都与地学密切相关，并且在地学中的应用也是特别广泛. 本章简要介绍分形理论在地质矿产资源勘探中的应用.

第一节　确定地球化学元素异常下限

地球化学元素异常下限是判别岩体是否成为矿体的标准. 地球化学异常下限的确定是勘察地球化学的一个基本问题，也是勘察地球化学应用于矿产勘察时决定成败的一个关键环节. 传统的确定地球化学元素异常下限的方法基于统计学理论，事先假设了数据是连续变化的.

自相似性（标度不变性）是地学中的一个普遍现象. 地壳中元素迁移聚集和成矿过程的混沌动力学机制，地质环境中非线性过程的相互作用是造成地壳元素含量与矿化不均匀分布的原因，并由此导致元素含量、矿床储量及其空间分布的分形结构. 矿产资源分布不均匀性是标度不变的. 品位分布的这种不均匀性有时可表现为极为完美的自相似性. 因此，可以使用分形理论来确定地球化学异常下限.

把地球化学元素的数据记为（x_i，y_i，z_i），其中 x_i，y_i 代表地理位置，z_i 代表元素含量. 我们把（x_i，y_i，z_i）构成的曲面称为含量曲面. 显然含量曲面应为具有分形性质的曲面，即采用不同含量尺度计算该曲面的有效面积（指元素含量高于给定的含量尺度的曲面部分的面积），曲面的有效面积在某标度区间上应具有标度不变性. 但是矿石和作为非矿石的岩石的化学元素含量曲面的有效面积应该具有不同的标度不变区间，即当元素含量尺度从异常下限以下升到异常下限以上时，化学元素含量曲面的有效面积应该具有明显不同的维数.

先将含量曲面的投影平面用矩形网络分割为边长为 $\Delta x\times\Delta y$ 的矩形，第 k 个矩形记为 $abcd$，这四个点的投影高度为 h_{ak}，h_{bk}，h_{ck}，h_{dk}（即四个点处元素的含量）. 选取含量尺度 r，当 h_{ak}，h_{bk}，h_{ck}，h_{dk} 均大于或等于 r 时，计算该投影网格对应的小曲面的有效面积，近似有效面积公式为

$$A_k(r) = \frac{1}{2}\Big[\sqrt{\Delta x^2 + (h_{ak} - h_{dk})^2} \times \sqrt{\Delta y^2 + (h_{dk} - h_{ck})^2}$$

$$+ \sqrt{\Delta x^2 + (h_{bk} - h_{ck})^2} \times \sqrt{\Delta y^2 + (h_{ak} - h_{bk})^2}\Big],$$

否则 $A_k(r)=0$.

整个投影网络对应在曲面上的覆盖有效总面积可近似为

$$A(r) = \sum_{k=1}^{N} A_k(r),$$

式中 N 是小矩形数目. 取不同的 r 值,将得到不同的 $A(r)$.

此分形模型为

$$A(r) = Cr^{-D},\ r > 0,$$

其中 r 表示含量尺度,C 称为比例常数,D 称为分维数,$A(r)$是尺度 r 下的投影面积.

为了求出分维数 D,将观测数据($A(r_1)$, $A(r_2)$, …, $A(r_n)$)和(r_1, r_2, …, r_n)绘在双对数坐标图上,如果其点大致分布在一条直线上的话,分维数 D 可以利用直线的斜率的绝对值来表示,也就是说,将观测数据($A(r_1)$, $A(r_2)$, …, $A(r_n)$)和(r_1, r_2, …, r_n)代入此分形模型,然后两边取对数,得线性回归模型

$$\ln A(r) = -D\ln r + \ln C.$$

用最小二乘法求出斜率 D 的估计量,即为分维数. 如果其点大致分布在两段直线上,可采用分段拟合. 为了提高分界点确定的客观性,在两个区间用最小二乘法回归时用最优化方法确定分界点. 其基本思想是,任取数据点 i_0,求出各区间拟合的直线与原始数据之间的剩余平方和 $S_i (i = 1,\ 2)$ 在两个区间的总和

$$\begin{aligned} S &= S_1 + S_2 \\ &= \sum_{i=1}^{i_0}[\ln A(r_i) + D_1 \ln r_i - \ln C_1]^2 \\ &\quad + \sum_{i=i_0+1}^{N}[\ln A(r_i) + D_2 \ln r_i - \ln C_2]^2. \end{aligned}$$

让 i_0 取遍所有数据点,使 S 取最小值的点 k 即为分界点,D_1 和 D_2 分别为相应区间的斜率分维数. 为了检验回归方程的显著性,对每个回归方程都要进行相关系数检验及方差分析检验.

用所求维数确定出异常下限后,用大于异常下限的数据绘制异常图.

采用某强烈切割区的 1∶20 万地球化学数据,对 Ag, Mo, Bi, Pb 四种元素使用上面的方法处理分析,得到各元素的异常下限值(表 7.4.1,图 7.4.1,图

7.4.2,图 7.4.3,图 7.4.4)和异常图形(图 7.4.5,图 7.4.6,图 7.4.7,图 7.4.8,图 7.4.9,图 7.4.10,图 7.4.11,图 7.4.12).

表 7.4.1　某地区的 1∶20 万地球化学数据

元素	异常下限	异常点数	第一段维数	第二段维数
Ag	130.452 545	246	0.7	2.3
Mo	1.001 997	367	0.6	4.6
Bi	0.501 207	253	0.7	0.4
Pb	35.085 5	275	0.6	2.9

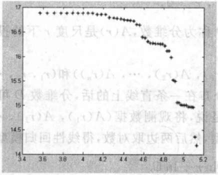


图 7.4.1　Ag 元素 $\ln A-\ln r$ 图

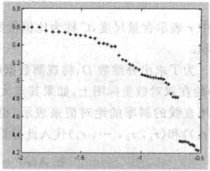


图 7.4.2　Bi 元素 $\ln A-\ln r$ 图

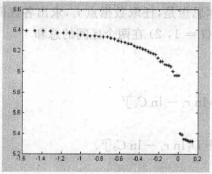


图 7.4.3　Mo 元素 $\ln A-\ln r$ 图

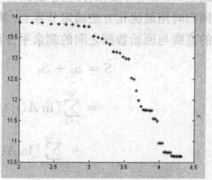


图 7.4.4　Pb 元素 $\ln A-\ln r$ 图

从图形上可以看出,根据这一方法对该地区的 Ag、Bi、Mo、Pb 四种地球化学元素数据的处理,区分了地球化学背景值和异常值,取得了较好的效果,表明该方法能提供独特的信息,说明分维数是划分地球化学异常的一个重要参数.

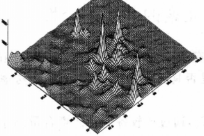


图 7.4.5　Ag 的原始数据三维图

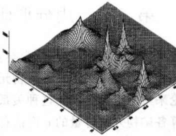


图 7.4.6　Ag 的三维异常图

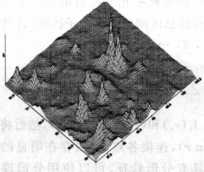


图 7.4.7　Bi 的原始数据三维图

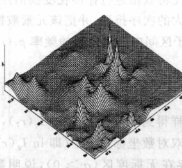


图 7.4.8　Bi 的三维异常图

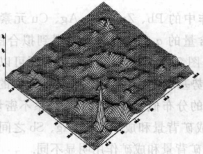


图 7.4.9　Mo 的原始数据三维图

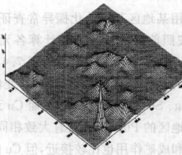


图 7.4.10　Mo 的三维异常图

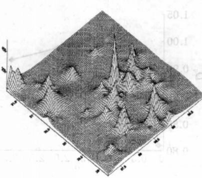


图 7.4.11　Pb 的原始数据三维图

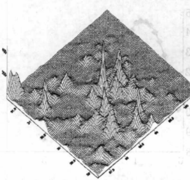


图 7.4.12　Pb 的三维异常图

第二节　用分形研究元素的共生组合性

在矿物的勘探和开采中，常常会发现有些矿物总是共生的．通常我们用统计分析中的相关分析来判别这些矿物元素是否具有相关性．同样，也可以考虑采用分形理论来考察矿物元素的这种关系．

计算各矿物元素含量的 q 次信息维 D_q，通过比较各矿物元素含量的 q 次信息维 D_q 的变化趋势来判定它们的关系是否密切．

首先将数据进行标准化变换消除量纲，再将标准化变换后的各元素数据按从小到大的次序排列，并把该元素数据分布的总区间分成 r 个子区间．计算进入第 i 个子区间内的随机数的频率 $p_i(i = 1, 2, \cdots, r)$，记

$$I_q(r) = \frac{1}{1-q}\sum_{i=1}^{r} p_i^q(r).$$

这样得到了数据 $I_q(r_1)$，$I_q(r_2)$，…，$I_q(r_n)$和$(r_1, r_2, \cdots, r_n)$，然后将该数据绘在双对数坐标系统中（即 $\ln I_q(r) - \ln r$），连接各点，曲线存在明显的直线段，则存在无标度区（$q \geqslant 0$），说明数据具有分形特征，可以使用分形理论来处理．

选用某地区某年的化探异常查证工作中的 Pb、Zn、Sb、Ag、Cu 元素含量数据，按照前面的计算方法计算各元素含量的 q 次信息维 D_q，得到拟合图（图 7.4.13，图 7.4.14，图 7.4.15，图 7.4.16，图 7.4.17）．从以上的拟合图可以看出该地区的 Pb，Zn 元素的信息维 D_q 变化趋势基本一致，说明了 Pb、Zn 元素的关系密切，Ag、Sb 之间的关系次之，而 Cu 元素的分布似乎与以上元素关系不密切．这说明该地区的 Pb、Zn 元素有大致相同的成矿背景和成矿作用，Ag、Sb 之间的成矿背景和成矿作用也比较接近，但 Cu 的成矿背景和成矿作用明显不同．

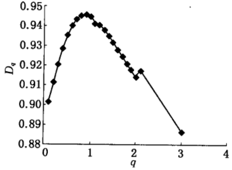


图 7.4.13　Pb 元素的 D_q 拟合图

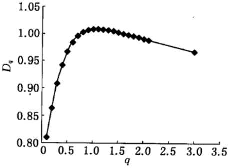


图 7.4.14　Zn 元素的 D_q 拟合图

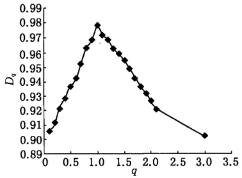


图 7.4.15 Sb 元素的 D_q 拟合图

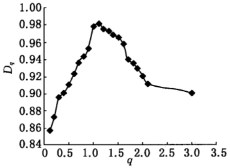


图 7.4.16 Ag 元素 D_q 拟合图

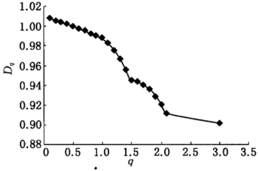


图 7.4.17 Cu 元素 D_q 拟合图

参考文献

1. 张济忠. 分形. 北京:清华大学出版社,1995

2. 孙霞,吴自强,黄昀. 分形原理及其应用. 合肥:中国科技大学出版社,2003

3. 李水根,吴纪桃. 分形与小波. 北京:科学出版社,2002

4. [法]B. 曼德尔布洛特. 分形对象(形、机遇与维数). 文志英,苏虹译. 北京:世界图书出版公司,1999

5. [英]肯尼思·法尔科内. 分形几何——数学基础及其应用. 曾文曲,等译. 沈阳:东北大学出版社,1991

6. [英]肯尼思·法尔科内. 分形几何中的技巧. 曾文曲,等译. 沈阳:东北大学出版社,1999

7. 申维. 分形混沌与矿产预测. 北京:地质出版社,2002

8. 李长江,麻士华. 矿产勘探中的分形、混沌与 ANN. 北京:地质出版社,1999

9. 王占元. 广义 M-J 集的分形机理. 大连:大连理工大学出版社,2002

10. 陈聆,魏友华,郭科. 用含量面积法确定深切割地区地球化学异常. 西南师范大学学报(自然科学版),2004, 29(5)

11. 郭科,施泽进,唐菊兴,等. 用多重分形研究元素的共生组合性. 电子科技大学学报,2004, 33(2)